=> d his

(FILE 'HOME' ENTERED AT 14:08:54 ON 26 MAR 2002)

(FILE 'MEDLINE, CABA, CAPLUS, BIOTECHNO, CONFSCI, EMBASE, BIOTECHDS, WPIDS' ENTERED AT 14:09:39 ON 26 MAR 2002
L1 2370 S MITCHELL W?/AU
L2 592 S STRATTON C?/AU
L3 32 S L1 AND L2
L4 40243 S ?CHLAMYDI?
L5 53 S L4 AND (L1 OR L2)
L6 55 S L3 OR L5
L7 30 DUP REM L6 (25 DUPLICATES REMOVED)

=> d ibib ab 17 1-30; fil hom

L7 ANSWER 1 OF 30 MEDLINE DUPLICATE 1
ACCESSION NUMBER: 2001371246 MEDLINE
DOCUMENT NUMBER: 21241020 PubMed ID: 11342681
TITLE: CSF oligoclonal bands in MS include antibodies against Chlamydophila antigens.
COMMENT: Comment in: Neurology. 2001 May 8;56(9):1126-7
Comment in: Neurology. 2001 May 8;56(9):1128-9
Comment in: Neurology. 2001 May 8;56(9):1130
AUTHOR: Yao S Y; **Stratton C W; Mitchell W M;** Sriram S
CORPORATE SOURCE: Department of Neurology, Vanderbilt University School of Medicine, Nashville, TN, USA.
SOURCE: NEUROLOGY, (2001 May 8) 56 (9) 1168-76.
Journal code: NZ0; 0401060. ISSN: 0028-3878.
PUB. COUNTRY: United States
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals
ENTRY MONTH: 200106
ENTRY DATE: Entered STN: 20010702
Last Updated on STN: 20010702
Entered Medline: 20010628

AB BACKGROUND: Considerable evidence suggests the role of an infectious agent in MS. The presence of Chlamydophila pneumoniae in CSF from patients with MS was shown earlier; to further examine this association the reactivity of the oligoclonal antibody response in the CSF of patients with MS to C pneumoniae antigens was determined and compared with other antigens. METHODS: Seventeen patients with MS and 14 control subjects with other neurologic disease were studied. Affinity-driven immunoblot studies and solid-phase adsorption of CSF oligoclonal bands by elementary body antigens of C pneumoniae, viral antigens (measles and herpes simplex virus-1), bacterial antigen (Escherichia coli and Staphylococcus aureus), and heat shock protein-60 were performed. RESULTS: Affinity-driven immunoblot studies demonstrated reactivity of oligoclonal bands in CSF samples from 16 patients with MS against elementary body antigens of C pneumoniae. None of the control subjects showed a prominent reactivity to elementary body antigens of C pneumoniae. In 14 of 17 patients with MS examined, oligoclonal bands were adsorbed either partially or completely from the CSF by elementary body antigens of C pneumoniae, but not by myelin basic protein, heat shock protein-60, or bacterial or viral antigens. In three patients with subacute sclerosing panencephalitis, adsorption of oligoclonal bands was seen with measles virus antigens but not with elementary body antigens of C pneumoniae. CONCLUSIONS: Oligoclonal bands in CSF of patients with MS include antibodies against Chlamydophila antigens.

L7 ANSWER 2 OF 30 MEDLINE DUPLICATE 2

ACCESSION NUMBER: 2001435272 MEDLINE
DOCUMENT NUMBER: 21229846 PubMed ID: 11331036
TITLE: Regulation by IFN-beta of inducible nitric oxide synthase and interleukin-12/p40 in murine macrophages cultured in the presence of **Chlamydia** pneumoniae antigens.
AUTHOR: Yao S Y; Ljunggren-Rose A; **Stratton C W; Mitchell W M**; Sriram S
CORPORATE SOURCE: Department of Neurology, Vanderbilt University School of Medicine, Nashville, TN 37212, USA.
SOURCE: JOURNAL OF INTERFERON AND CYTOKINE RESEARCH, (2001 Mar) 21 (3) 137-46.
Journal code: CD4; 9507088. ISSN: 1079-9907.
PUB. COUNTRY: United States
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200108
ENTRY DATE: Entered STN: 20010806
Last Updated on STN: 20010806
Entered Medline: 20010802

AB **Chlamydia** pneumoniae has been demonstrated in the cerebrospinal fluid (CSF) of patients with multiple sclerosis (MS). Interferon-beta (IFN-beta) has favorable effects on the clinical course of MS. We investigated whether the beneficial effects of IFN-beta in MS may involve its role in regulating nitric oxide (NO) and interleukin-12 (IL-12) in macrophages, as these immune modulators form part of the innate immune response to intracellular pathogens, such as C. pneumoniae. Murine macrophages in cultures exposed to elementary body antigens or recombinant major outer membrane protein (rMOMP) of C. pneumoniae demonstrate a significant increase in NO as well as production of IL-12/p40 in culture supernatants compared with basal levels. Addition of murine IFN-beta increased NO activity in murine macrophages cultured with **chlamydial** antigens. Addition of neutralizing anti-IFN-beta antibody prevented the NO increase. In contrast to its effect on inducible NO synthase (iNOS), IFN-beta reduced induction of IL-12/p40 following culture with either elementary body antigens or rMOMP. Inhibition was reversed with anti-IFN-beta antibody. If C. pneumoniae infection is responsible for the inflammatory response in the pathogenesis of MS, the beneficial effects of IFN-beta in MS may be due to its enhancing intracellular NO activity while inhibiting secretion of the proinflammatory cytokine, IL-12.

L7 ANSWER 3 OF 30 MEDLINE DUPLICATE 5
ACCESSION NUMBER: 2000111212 MEDLINE
DOCUMENT NUMBER: 20111212 PubMed ID: 10642692
TITLE: Pyoderma gangrenosum and **Chlamydia** pneumoniae infection in a diabetic man: pathogenic role or coincidence?.
AUTHOR: Vannucci S A; **Mitchell W M; Stratton C W** ; King L E Jr
CORPORATE SOURCE: Department of Medicine, Division of Dermatology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-5227, USA.
SOURCE: JOURNAL OF THE AMERICAN ACADEMY OF DERMATOLOGY, (2000 Feb) 42 (2 Pt 1) 295-7.
Journal code: HVG; 7907132. ISSN: 0190-9622.
PUB. COUNTRY: United States
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200002
ENTRY DATE: Entered STN: 20000309
Last Updated on STN: 20000309

Entered Medline: 20000224

AB **Chlamydia** Pneumoniae is not a known cause of skin infections, but unusual pathogens cause chronic infections in diabetic patients. Multiple idiopathic pyoderma gangrenosum-like (PG-like) lesions were refractory to multiple therapeutic agents in a diabetic patient who had C pneumoniae identified by serologic tests and polymerase chain reaction. Based on complete resolution by prolonged anti-**chlamydial** antibiotic therapy and concomitant decrease in serologic and titers determined by polymerase chain reactions, the PG-like lesions were presumed to be due to C pneumoniae.

L7 ANSWER 4 OF 30 MEDLINE DUPLICATE 6
ACCESSION NUMBER: 1999206606 MEDLINE
DOCUMENT NUMBER: 99206606 PubMed ID: 10192388
TITLE: Comparative genomes of **Chlamydia** pneumoniae and C. trachomatis.
AUTHOR: Kalman S; **Mitchell W**; Marathe R; Lammel C; Fan J; Hyman R W; Olinger L; Grimwood J; Davis R W; Stephens R S
CORPORATE SOURCE: Stanford DNA Sequencing and Technology Center, Stanford University, California 94305, USA.
SOURCE: NATURE GENETICS, (1999 Apr) 21 (4) 385-9. Journal code: BRO; 9216904. ISSN: 1061-4036.
PUB. COUNTRY: United States
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-AE001273; GENBANK-AE001363
ENTRY MONTH: 199904
ENTRY DATE: Entered STN: 19990511
Last Updated on STN: 19990511
Entered Medline: 19990426

AB **Chlamydia** are obligate intracellular eubacteria that are phylogenetically separated from other bacterial divisions. C. trachomatis and C. pneumoniae are both pathogens of humans but differ in their tissue tropism and spectrum of diseases. C. pneumoniae is a newly recognized species of **Chlamydia** that is a natural pathogen of humans, and causes pneumonia and bronchitis. In the United States, approximately 10% of pneumonia cases and 5% of bronchitis cases are attributed to C. pneumoniae infection. Chronic disease may result following respiratory-acquired infection, such as reactive airway disease, adult-onset asthma and potentially lung cancer. In addition, C. pneumoniae infection has been associated with atherosclerosis. C. trachomatis infection causes trachoma, an ocular infection that leads to blindness, and sexually transmitted diseases such as pelvic inflammatory disease, chronic pelvic pain, ectopic pregnancy and epididymitis. Although relatively little is known about C. trachomatis biology, even less is known concerning C. pneumoniae. Comparison of the C. pneumoniae genome with the C. trachomatis genome will provide an understanding of the common biological processes required for infection and survival in mammalian cells. Genomic differences are implicated in the unique properties that differentiate the two species in disease spectrum. Analysis of the 1,230,230-nt C. pneumoniae genome revealed 214 protein-coding sequences not found in C. trachomatis, most without homologues to other known sequences. Prominent comparative findings include expansion of a novel family of 21 sequence-variant outer-membrane proteins, conservation of a type-III secretion virulence system, three serine/threonine protein kinases and a pair of parologous phospholipase-D-like proteins, additional purine and biotin biosynthetic capability, a homologue for aromatic amino acid (tryptophan) hydroxylase and the loss of tryptophan biosynthesis genes.

L7 ANSWER 5 OF 30 MEDLINE DUPLICATE 7
ACCESSION NUMBER: 1999328202 MEDLINE

DOCUMENT NUMBER: 99328202 PubMed ID: 10401775
TITLE: **Chlamydia** pneumoniae infection of the central nervous system in multiple sclerosis.
COMMENT: Comment in: Ann Neurol. 1999 Jul;46(1):4-5
Comment in: Ann Neurol. 2000 Mar;47(3):408-9; discussion 409-11
Comment in: Ann Neurol. 2000 Mar;47(3):408; discussion 409-11
Comment in: Ann Neurol. 2000 Sep;48(3):399
Comment in: Ann Neurol. 2000 Sep;48(3):399-400
Comment in: Ann Neurol. 2000 Sep;48(3):400
Comment in: Ann Neurol. 2001 Jan;49(1):135
AUTHOR: Sriram S; **Stratton C W**; Yao S; Tharp A; Ding L; Bannan J D; **Mitchell W M**
CORPORATE SOURCE: Department of Neurology, Vanderbilt School of Medicine, Nashville, TN, USA.
SOURCE: ANNALS OF NEUROLOGY, (1999 Jul) 46 (1) 6-14.
Journal code: 7707449. ISSN: 0364-5134.
PUB. COUNTRY: United States
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199908
ENTRY DATE: Entered STN: 19990816
Last Updated on STN: 20020314
Entered Medline: 19990802

AB Our identification of **Chlamydia** pneumoniae in the cerebrospinal fluid (CSF) of a patient with multiple sclerosis (MS) led us to examine the incidence of this organism in the CSF from 17 patients with relapsing-remitting MS, 20 patients with progressive MS, and 27 patients with other neurological diseases (OND). CSF samples were examined for C pneumoniae by culture, polymerase chain reaction assays, and CSF immunoglobulin (Ig) reactivity with C pneumoniae elementary body antigens. C pneumoniae was isolated from CSF in 64% of MS patients versus 11% of OND controls. Polymerase chain reaction assays demonstrated the presence of C pneumoniae MOMP gene in the CSF of 97% of MS patients versus 18% of OND controls. Finally, 86% of MS patients had increased CSF antibodies to C pneumoniae elementary body antigens as shown by enzyme-linked immunosorbent assay absorbance values that were 3 SD greater than those seen in OND controls. The specificity of this antibody response was confirmed by western blot assays of the CSF, using elementary body antigens. Moreover, CSF isoelectric focusing followed by western blot assays revealed cationic antibodies against C pneumoniae. Infection of the central nervous system with C pneumoniae is a frequent occurrence in MS patients. Although the organism could represent the pathogenetic agent of MS, it may simply represent a secondary infection of damaged central nervous system tissue. A therapeutic trial directed at eliminating C pneumoniae from the central nervous system may provide additional information on its role in MS.

L7 ANSWER 6 OF 30 MEDLINE DUPLICATE 10
ACCESSION NUMBER: 1999000809 MEDLINE
DOCUMENT NUMBER: 99000809 PubMed ID: 9784136
TITLE: Genome sequence of an obligate intracellular pathogen of humans: **Chlamydia** trachomatis.
COMMENT: Comment in: Science. 1998 Oct 23;282(5389):638-9
AUTHOR: Stephens R S; Kalman S; Lammel C; Fan J; Marathe R; Aravind L; **Mitchell W**; Olinger L; Tatusov R L; Zhao Q; Koonin E V; Davis R W
CORPORATE SOURCE: Program in Infectious Diseases, University of California, Berkeley, CA 94720, USA.. ctgenome@socrates.berkeley.edu
CONTRACT NUMBER: AI 39258 (NIAID)
SOURCE: SCIENCE, (1998 Oct 23) 282 (5389) 754-9.

Journal code: UJ7; 0404511. ISSN: 0036-8075.
PUB. COUNTRY: United States
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-AE001275; GENBANK-AE001276; GENBANK-AE001277; GENBANK-AE001278; GENBANK-AE001279; GENBANK-AE001280; GENBANK-AE001281; GENBANK-AE001282; GENBANK-AE001283; GENBANK-AE001284; GENBANK-AE001285; GENBANK-AE001286; GENBANK-AE001287; GENBANK-AE001288; GENBANK-AE001289; GENBANK-AE001290; GENBANK-AE001291; GENBANK-AE001292; GENBANK-AE001293; GENBANK-AE001294; GENBANK-AE001295; GENBANK-AE001296; GENBANK-AE001297; GENBANK-AE001298; GENBANK-AE001299; GENBANK-AE001300; GENBANK-AE001301; GENBANK-AE001302; GENBANK-AE001303; GENBANK-AE001304
ENTRY MONTH: 199811
ENTRY DATE: Entered STN: 19990106
Last Updated on STN: 20000303
Entered Medline: 19981109

AB Analysis of the 1,042,519-base pair **Chlamydia** trachomatis genome revealed unexpected features related to the complex biology of **chlamydiae**. Although **chlamydiae** lack many biosynthetic capabilities, they retain functions for performing key steps and interconversions of metabolites obtained from their mammalian host cells. Numerous potential virulence-associated proteins also were characterized. Several eukaryotic chromatin-associated domain proteins were identified, suggesting a eukaryotic-like mechanism for **chlamydial** nucleoid condensation and decondensation. The phylogenetic mosaic of **chlamydial** genes, including a large number of genes with phylogenetic origins from eukaryotes, implies a complex evolution for adaptation to obligate intracellular parasitism.

L7 ANSWER 7 OF 30 MEDLINE DUPLICATE 11
ACCESSION NUMBER: 1998145402 MEDLINE
DOCUMENT NUMBER: 98145402 PubMed ID: 9484408
TITLE: Multiple sclerosis associated with **Chlamydia** pneumoniae infection of the CNS.
COMMENT: Comment in: Neurology. 2001 Aug 28;57(4):746
AUTHOR: Sriram S; **Mitchell W; Stratton C**
CORPORATE SOURCE: Department of Neurology, Vanderbilt University Medical Center, Nashville, TN 37212, USA.
SOURCE: NEUROLOGY, (1998 Feb) 50 (2) 571-2.
Journal code: 0401060. ISSN: 0028-3878.
PUB. COUNTRY: United States
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals
OTHER SOURCE: GENBANK-AF131889
ENTRY MONTH: 199803
ENTRY DATE: Entered STN: 19980326
Last Updated on STN: 20020130
Entered Medline: 19980316

L7 ANSWER 8 OF 30 MEDLINE DUPLICATE 13
ACCESSION NUMBER: 92346653 MEDLINE
DOCUMENT NUMBER: 92346653 PubMed ID: 1638577
TITLE: Fluoroquinolone antibiotics: properties of the class and individual agents.
AUTHOR: **Stratton C**
CORPORATE SOURCE: Department of Pathology, Vanderbilt University Medical Center, Nashville, Tennessee.
SOURCE: CLINICAL THERAPEUTICS, (1992 May-Jun) 14 (3) 348-75; discussion 347. Ref: 121

Journal code: CPE; 7706726. ISSN: 0149-2918.
PUB. COUNTRY: United States
Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199209
ENTRY DATE: Entered STN: 19920911
Last Updated on STN: 19920911
Entered Medline: 19920901

AB The broad spectrum of activity and bactericidal nature of the fluoroquinolones, together with their excellent absorption, rapid distribution, and high tissue concentration, make them excellent therapeutic agents for the management of a number of complicated community-acquired and nosocomial infections of the urinary tract, bone and soft tissue, gastrointestinal tract, and prostate, as well as some respiratory tract infections and sexually transmitted diseases. Data are presented and reviewed concerning the in vitro activity, pharmacology, and clinical use of ciprofloxacin, norfloxacin, and ofloxacin, which have been available for some time, and lomefloxacin and temafloxacin, which are recently approved agents. The comparable qualities and differences in activity and clinical applications of these agents are considered. For many infections in selected patients, quinolones are excellent substitutes for parenteral agents. In general, adverse effects have been infrequent and rarely require drug discontinuation. Significant interactions, such as with theophylline and caffeine, have occurred but are quinolone dependent. Antacids can markedly impair the absorption of all quinolones. Because emerging resistance to Pseudomonas and Staphylococcus species have been observed, the improper use of the quinolones must be avoided, and the clinician must be aware that an unfavorable response may signal resistance. The development of future agents with better gram-positive activity, improved gram-negative coverage, and activity against unusual pathogens such as **Chlamydia** species and Mycobacterium species, will make these oral agents invaluable. Assessing the usefulness and safety of these antibiotics in children is an ongoing challenge.

L7 ANSWER 9 OF 30 MEDLINE
ACCESSION NUMBER: 2001314629 MEDLINE
DOCUMENT NUMBER: 21244261 PubMed ID: 11346358
TITLE: Diabetic foot ulcers and **Chlamydia** pneumoniae: innocent bystander or opportunistic pathogen?.
AUTHOR: King L E Jr; Bushman T; **Stratton C W**; **Mitchell W M**
SOURCE: ARCHIVES OF DERMATOLOGY, (2001 May) 137 (5) 671-2.
Journal code: 6WU; 0372433. ISSN: 0003-987X.
PUB. COUNTRY: United States
Letter
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals
ENTRY MONTH: 200106
ENTRY DATE: Entered STN: 20010611
Last Updated on STN: 20010611
Entered Medline: 20010607

L7 ANSWER 10 OF 30 MEDLINE
ACCESSION NUMBER: 2000092443 MEDLINE
DOCUMENT NUMBER: 20092443 PubMed ID: 10628821
TITLE: Does **Chlamydia** pneumoniae play a role in the pathogenesis of multiple sclerosis?.
AUTHOR: **Stratton C W**; **Mitchell W M**; Sriram S
SOURCE: JOURNAL OF MEDICAL MICROBIOLOGY, (2000 Jan) 49 (1) 1-3.
Journal code: J2N; 0224131. ISSN: 0022-2615.

PUB. COUNTRY: ENGLAND: United Kingdom
Editorial
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200001
ENTRY DATE: Entered STN: 20000124
Last Updated on STN: 20000124
Entered Medline: 20000113

L7 ANSWER 11 OF 30 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 3
ACCESSION NUMBER: 2000:688466 CAPLUS
DOCUMENT NUMBER: 133:249334
TITLE: Methods and reagents for the diagnosis and treatment of multiple sclerosis caused by **Chlamydia**
INVENTOR(S): **Stratton, Charles W.; Mitchell, William M.**; Yao, Song-yi; Bannan, Jason D.; Ljunggren-Rose, Asa; Sriram, Subramaniam
PATENT ASSIGNEE(S): Vanderbilt University, USA
SOURCE: PCT Int. Appl., 102 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000057187	A2	20000928	WO 2000-US7226	20000317
WO 2000057187	A3	20010419		
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
EP 1166117	A2	20020102	EP 2000-916513	20000317
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				

PRIORITY APPLN. INFO.:
US 1999-125598P P 19990319
US 2000-176662P P 20000118
US 2000-176784P P 20000118
US 2000-176940P P 20000118
WO 2000-US7226 W 20000317

AB The invention features methods and reagents for the diagnosis, monitoring, and treatment of multiple sclerosis. The invention is based in part on the discovery that **Chlamydia** is present in patients with multiple sclerosis, and that anti-**chlamydial** agents improve or sustain neurol. function in these patients.

L7 ANSWER 12 OF 30 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 4
ACCESSION NUMBER: 2000:335519 CAPLUS
DOCUMENT NUMBER: 133:1493
TITLE: **Chlamydia** pneumoniae genome sequence
INVENTOR(S): Stephens, Richard; **Mitchell, Wayne**; Kalman, Sue; Davis, Ronald
PATENT ASSIGNEE(S): The Regents of the University of California, USA
SOURCE: PCT Int. Appl., 330 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English

Searched by Barb O'Bryen, STIC 308-4291

FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000027994	A2	20000518	WO 1999-US26923	19991112
WO 2000027994	A3	20001123		

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

AU 2000017223	A5	20000529	AU 2000-17223	19991112
EP 1133572	A2	20010919	EP 1999-960323	19991112

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO

PRIORITY APPLN. INFO.:

US 1998-108279P	P	19981112
US 1999-128606P	P	19990408
WO 1999-US26923	W	19991112

AB The **Chlamydia** pneumoniae genome sequence and anal. of the encoded polypeptides and RNAs are provided. The C. pneumoniae genome contains 187,711 addnl. nucleotides compared to the C. trachomatis genome, and the 214 coding sequences not found in C. trachomatis account for most of the increased genome size. The majority of these addnl. genes lack identifiable homologs to genes from other organisms, and probably are essential for specific attributes that define the differential biol., tropism, and pathogenesis of C. trachomatis and C. pneumoniae. The C. pneumoniae gene nucleic acid compns. find use in identifying homologous or related proteins and the DNA sequences encoding such proteins; in producing compns. that modulate the expression or function of the protein; and in studying assocd. physiol. pathways. In addn., modulation of the gene activity in vivo is used for prophylactic and therapeutic purposes, such as identification of cell type based on expression, and the like.

L7 ANSWER 13 OF 30 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 8
ACCESSION NUMBER: 1998:752291 CAPLUS
DOCUMENT NUMBER: 130:10609
TITLE: Diagnosis and management of infection caused by **Chlamydia**
INVENTOR(S): **Mitchell, William M.; Stratton, Charles W.**
PATENT ASSIGNEE(S): Vanderbilt University, USA
SOURCE: PCT Int. Appl., 139 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 4
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9850074	A2	19981112	WO 1998-US9237	19980506
WO 9850074	A3	19990819		

W: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES,

FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

US 2001002421	A1	20010531	US 1998-25176		19980218
US 6258532	B1	20010710			
AU 9872899	A1	19981127	AU 1998-72899		19980506
EP 981372	A2	20000301	EP 1998-920292		19980506
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI					
ZA 9803798	A	20000307	ZA 1998-3798		19980506

PRIORITY APPLN. INFO.:

US 1997-45689P	P	19970506
US 1997-45739P	P	19970506
US 1997-45779P	P	19970506
US 1997-45780P	P	19970506
US 1997-45784P	P	19970506
US 1997-45787P	P	19970506
US 1997-911593	A	19970814
US 1998-25176	A2	19980218
US 1998-25521	A2	19980218
US 1998-25174	A	19980218
WO 1998-US9237	W	19980506

AB A combination of agents directed toward various stages of the **chlamydial** life cycle is effective in substantially reducing infection. These include agents targeted against the cryptic phase (e.g. nitroarom. compds.), elementary body phase (e.g. disulfide reducing agents), and replicating phase, probenecid, and antiporphyric agents. **Chlamydia**-free cell lines and animals can be obtained, and **Chlamydia** infections can be treated, by use of .gtoreq.2 such agents. **Chlamydia** infections may be diagnosed or monitored by immunoassays (e.g. ELISA or antigen capture assay) for the cysteine-rich major outer membrane protein or for specific antigenic peptides, DNA amplification assays (e.g. PCR) for **chlamydial** genes, and Western blot assays. Thus, a multiple sclerosis patient showing progressive limb impairment was diagnosed with C. pneumoniae infection by cerebrospinal fluid PCR and culture; treatment with rifampin (300 mg twice a day for 2 mo against the elementary body/reticulate body transition), flagyl (500 mg twice a day for 5 mo against the stationary phase reticulate body), and ofloxacin (for 2 mo) and Bactrim (double strength twice a day) and levaquin (500 mg/day) for 5 mo against the replicating reticulate body resulted in marked improvement in all aspects of neurol. function and an ability to return to work and routine athletic activities.

L7 ANSWER 14 OF 30 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 9

ACCESSION NUMBER: 1998:124043 CAPLUS
DOCUMENT NUMBER: 128:201045
TITLE: Compositions of **antichlamydial** agents for the diagnosis and management of infection caused by **chlamydia**
INVENTOR(S): **Mitchell, William M.; Stratton, Charles W.**
PATENT ASSIGNEE(S): Vanderbilt University, USA; Mitchell, William M.; Stratton, Charles W.
SOURCE: PCT Int. Appl., 83 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 4
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9806435	A2	19980219	WO 1997-US14402	19970814
WO 9806435	A3	19980409		
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE,				

DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ,
LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL,
PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US,
UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA,
GN, ML, MR, NE, SN, TD, TG

AU 9741516 A1 19980306 AU 1997-41516 19970814
PRIORITY APPLN. INFO.: US 1996-23921P P 19960814
WO 1997-US14402 W 19970814

AB The invention provides a unique approach for the diagnosis and management of infections by **Chlamydia** species, particularly C. pneumoniae. The invention is based, in part, on the discovery that a combination of agents directed toward the various stages of the **chlamydial** life cycle is effective in substantially reducing infection. Products comprising combination of **antichlamydial** agents, compns., and pharmaceutical packs are also described.

L7 ANSWER 15 OF 30 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2002:65982 CAPLUS
DOCUMENT NUMBER: 136:133602
TITLE: Identification of antigenic peptide sequences
INVENTOR(S): **Mitchell, William M.; Stratton, Charles W.**
PATENT ASSIGNEE(S): Vanderbilt University, USA
SOURCE: U.S., 40 pp., Cont.-in-part of U.S. Ser. No. 911,593, abandoned.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 4
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6340463	B1	20020122	US 1998-25596	19980218

PRIORITY APPLN. INFO.: US 1996-23921P P 19960814
US 1997-911593 B2 19970814

AB Identification of linear amino acid antigenic sequences for the prodn. of both polyclonal and monoclonal antibodies to defined antigenic domains is described. Also described are antigenic peptides identified by the described methods and antibodies thereto.
REFERENCE COUNT: 7 THERE ARE 7 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 16 OF 30 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:397834 CAPLUS
DOCUMENT NUMBER: 135:2559
TITLE: Methods for in vitro susceptibility testing of **Chlamydia**
INVENTOR(S): **Stratton, Charles W.; Mitchell, William M.**
PATENT ASSIGNEE(S): USA
SOURCE: U.S. Pat. Appl. Publ., 9 pp., Cont.-in-part of U.S. Ser. No. 911,593, abandoned.
CODEN: USXXCO
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 4
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE

US 2001002421	A1	20010531	US 1998-25176	19980218
US 6258532	B1	20010710		
US 2002009802	A1	20020124	US 1998-25174	19980218
WO 9850074	A2	19981112	WO 1998-US9237	19980506
WO 9850074	A3	19990819		

W: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

AU 9872899	A1	19981127	AU 1998-72899	19980506
EP 981372	A2	20000301	EP 1998-920292	19980506

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI

PRIORITY APPLN. INFO.:

US 1997-911593	B2	19970814
US 1997-45689P	P	19970506
US 1997-45739P	P	19970506
US 1997-45779P	P	19970506
US 1997-45780P	P	19970506
US 1997-45784P	P	19970506
US 1997-45787P	P	19970506
US 1998-25174	A	19980218
US 1998-25176	A2	19980218
US 1998-25521	A2	19980218
WO 1998-US9237	W	19980506

AB Methods for detg. the susceptibility of intracellular pathogens, particularly **Chlamydia**, to single or combination of test agents are described. The methods can be used for in vitro or in vivo evaluation of agents that can be used as therapeutic agents in the treatment/eradication of pathogen infection in general or to target a specific infected organ. Assays which utilize nucleic amplification techniques (e.g., PCR) to det. effectiveness of the agent(s) evaluated are also described.

L7 ANSWER 17 OF 30 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2001:607994 CAPLUS

DOCUMENT NUMBER: 136:84535

TITLE: Association of multiple sclerosis with **Chlamydia** pneumoniae: Demonstration of the 16S rRNA gene and immunoreactivity of CSF cationic antibodies against C. pneumoniae antigens

AUTHOR(S): Sriram, S.; **Stratton, C. W.**; Yao, S.; Bannan, J. D.; **Mitchell, W. M.**

CORPORATE SOURCE: Department of Neurology, Vanderbilt School of Medicine, Nashville, TN, USA

SOURCE: Genes and Viruses in Multiple Sclerosis (2001), 221-229. Editor(s): Hommes, Otto R.; Clanet, Michel; Wekerle, Hartmut. Elsevier Science B.V.: Amsterdam, Neth.
CODEN: 69BRQ7

DOCUMENT TYPE: Conference

LANGUAGE: English

AB The immunoreactivity of the cationic antibodies in the CSF of multiple sclerosis (MS) patients was evaluated using affinity-driven immunoblot assays. Among the 17 CSF samples from MS patients, 88% contained DNA specific for the 16S rRNA gene of **Chlamydia** pneumoniae. MS patients had an increased Igs in the CSF, and part of the increase was represented as oligoclonal bands on isoelec. focusing gels. The development of an intrathecal immune response to C. pneumoniae was found to be a common occurrence in patients with MS. The cationic anti-EB

antibodies were present in patients with MS and might represent in part the specificity of the oligoclonal bands. The results of IEF/affinity-driven assays showed that a CNS immune response specifically to C. pneumoniae EB antigens was present in patients with MS.
REFERENCE COUNT: 19 THERE ARE 19 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 18 OF 30 CONFSCI COPYRIGHT 2002 CSA
ACCESSION NUMBER: 85:2268 CONFSCI
DOCUMENT NUMBER: 85002357
TITLE: Clinical evaluation of the MicroTrak direct specimen test for the detection of **Chlamydia** trachomatis in genital specimens
AUTHOR: Peters, T.H.; Lefkowitz, L.B.; Ratner, H.B.; Davidson, T.L.; **Stratton, C.W.**
CORPORATE SOURCE: Vanderbilt Univ. Med. Cent. and Metropolitan Health Dep., Nashville-Davidson County, Nashville, TN, USA
SOURCE: 1985, Abstracts available: American Society for Microbiology, Publication Department,1913 I St. NW, Washington, DC 20006, USA, Paper No. C 382.
Meeting Info.: 851 5000: American Society for Microbiology, 85th Annual Meeting (8515000). Las Vegas, NV (USA). 3-7 Mar 85. American Society for Microbiology.
DOCUMENT TYPE: Conference
FILE SEGMENT: DCCP
LANGUAGE: UNAVAILABLE

L7 ANSWER 19 OF 30 CONFSCI COPYRIGHT 2002 CSA
ACCESSION NUMBER: 86:15890 CONFSCI
DOCUMENT NUMBER: 86043671
TITLE: **Chlamydia** screening of male patients with gonorrhea
AUTHOR: Lefkowitz, L.B.; Peters, T.H.; Ratner, H.B.; Davidson, T.L.; **Stratton, C.W.**
CORPORATE SOURCE: Vanderbilt Univ., Nashville, TN, USA
SOURCE: American Society for Microbiology, 1913 I Street, N.W., Washington, DC 20006 (USA), Poster Paper No. C30.
Meeting Info.: 861 0146: American Society for Microbiology 86th Annual Meeting (8610146). Washington, DC (USA). 23-28 Mar 1986. American Society for Microbiology (ASM).
DOCUMENT TYPE: Conference
FILE SEGMENT: DCCP
LANGUAGE: UNAVAILABLE

L7 ANSWER 20 OF 30 CONFSCI COPYRIGHT 2002 CSA
ACCESSION NUMBER: 84:751 CONFSCI
DOCUMENT NUMBER: 84006418
TITLE: Comparison of immunoperoxidase staining with iodine staining for the detection of **Chlamydia** trachomatis inclusion bodies in McCoy cells
AUTHOR: **Stratton, C.W.**; Judson, F.N.; Simpson, E.B.; Jones, M.R.; Kasselberg, A.G.
CORPORATE SOURCE: Vanderbilt Univ. Hosp., Nashville, TN
SOURCE: Abstracts available: American Society for Microbiology, Publications Department, 1913 I St. NW, Washington, DC 20006, USA, Paper No. C105.
Meeting Info.: 841 0195: American Society for Microbiology 84th Annual Meeting (8410195). St. Louis, MO (USA). 4-9 Mar 84. American Society for Microbiology (ASM).
DOCUMENT TYPE: Conference
FILE SEGMENT: DCCP
LANGUAGE: UNAVAILABLE

L7 ANSWER 21 OF 30 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 2002091023 EMBASE
TITLE: Pharmacokinetic, pharmacodynamic and tolerability profiles of telithromycin, the first ketolide antimicrobial agent.
AUTHOR: **Stratton C.W.**
CORPORATE SOURCE: Dr. C.W. Stratton, Clinical Microbiology Laboratory, Vanderbilt Clinic, 21st and Edgehill, Nashville, TN 37232, United States. Charles.Stratton@mcmail.Vanderbilt.edu
SOURCE: Today's Therapeutic Trends, (2002) 20/1 (37-58).
Refs: 102
ISSN: 0741-2320 CODEN: TTTRDH
COUNTRY: United States
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 015 Chest Diseases, Thoracic Surgery and Tuberculosis
017 Public Health, Social Medicine and Epidemiology
030 Pharmacology
037 Drug Literature Index
038 Adverse Reactions Titles
LANGUAGE: English
SUMMARY LANGUAGE: English

AB The ketolides are the latest structural derivatives of erythromycin to be added to the macrolide family, with telithromycin being the first ketolide to undergo clinical evaluation. Telithromycin is a broad-spectrum and highly potent antimicrobial agent, which has a number of unique pharmacokinetic and pharmacodynamic properties that assure its clinical utility despite the increasing prevalence of macrolide resistance among the major respiratory tract pathogens. It appears that the most important factors in terms of structure-activity are the lack of the neutral sugar cladinose in position C3 as well as a C11/C12 carbamate group, which together markedly increase the affinity of telithromycin for its microbial target, the 23S ribosomal drug-binding pocket. This increased affinity is seen even in macrolide-resistant strains, and also results in concentration-dependent bactericidal activity and a prolonged post-antimicrobial effect against important respiratory tract pathogens. The microbiological spectrum of activity for telithromycin includes Streptococcus pneumoniae, Streptococcus pyogenes, Haemophilus influenzae, Moraxella catarrhalis, Legionella species, Mycoplasma pneumoniae, and **Chlamydia** pneumoniae, suggesting that telithromycin will play an important clinical role in the empirical treatment of community-acquired respiratory tract infections. The pharmacokinetic profile of telithromycin demonstrates that this drug can be administered once daily without regard for meals, and requires no dosage reduction in elderly patients or those with hepatic impairment. Telithromycin is well absorbed after oral administration, with rapid penetration into respiratory tissues and fluids, and as well is highly concentrated within white blood cells. Integration of pharmacokinetic and pharmacodynamic properties reveals that telithromycin has a high AUC/MIC ratio compared with macrolide antimicrobial agents, resulting in enhanced efficacy. Finally, telithromycin is well tolerated and has a low propensity for drug interactions. In summary, telithromycin promises to be a potent agent for the treatment of community-acquired respiratory tract infections.

L7 ANSWER 22 OF 30 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 2001185465 EMBASE
TITLE: Diabetic foot ulcers and **chlamydia** pneumoniae: Innocent bystander or opportunistic pathogen? [10].
AUTHOR: King L.E. Jr.; Bushman T.; **Stratton C.W.**; **Mitchell W.M.**
CORPORATE SOURCE: Dr. L.E. King Jr., Division of Dermatology, Vanderbilt Univ. Sch. of Med. 3983, Vanderbilt Clinic, Nashville, TN 37232-2556, United States. lloyd.king@amcmail.vanderbilt.edu
SOURCE: Archives of Dermatology, (2001) 137/5 (671-672).

Refs: 6
ISSN: 0003-987X CODEN: ARDEAC
COUNTRY: United States
DOCUMENT TYPE: Journal; Letter
FILE SEGMENT: 004 Microbiology
006 Internal Medicine
013 Dermatology and Venereology
LANGUAGE: English

L7 ANSWER 23 OF 30 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 2001241578 EMBASE
TITLE: Association of **Chlamydia** pneumoniae with chronic human diseases.
AUTHOR: **Stratton C.W.**
CORPORATE SOURCE: Dr. C.W. Stratton, Vanderbilt Univ. School of Medicine, Nashville, TN 37232, United States
SOURCE: Antimicrobics and Infectious Diseases Newsletter, (2000) 18/7 (49-55).
Refs: 141
ISSN: 1069-417X CODEN: AIDIEX
COUNTRY: United States
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 004 Microbiology
008 Neurology and Neurosurgery
011 Otorhinolaryngology
015 Chest Diseases, Thoracic Surgery and Tuberculosis
017 Public Health, Social Medicine and Epidemiology
LANGUAGE: English
SUMMARY LANGUAGE: English
AB It is apparent from this review that C. pneumoniae has been implicated in many chronic diseases of humans. Whether the role is that of innocent bystander, cause, or perhaps something in between remains to be determined. Regardless of the role of C. pneumoniae in these or other chronic diseases, this microorganism is becoming a major health concern. Considerable resources will be needed to determine its role in human disease. If C. pneumoniae proves to play an important role in any or all of these chronic diseases, its eventual control or eradication may do much to improve the health of countless persons.

L7 ANSWER 24 OF 30 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 2000017568 EMBASE
TITLE: Does **Chlamydia** pneumoniae play a role in the pathogenesis of multiple sclerosis?.
AUTHOR: **Stratton C.W.**; **Mitchell W.M.**; Sriram S.
CORPORATE SOURCE: C.W. Stratton, Department of Pathology, Vanderbilt School of Medicine, Nashville, TN 37232, United States. Charles.Stratton@.mcmail.Vanderbilt.edu
SOURCE: Journal of Medical Microbiology, (2000) 49/1 (1-3).
Refs: 28
ISSN: 0022-2615 CODEN: JMMIAV
COUNTRY: United Kingdom
DOCUMENT TYPE: Journal; Editorial
FILE SEGMENT: 004 Microbiology
008 Neurology and Neurosurgery
LANGUAGE: English

L7 ANSWER 25 OF 30 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 1999287545 EMBASE
TITLE: The role of **Chlamydia** in connective tissues diseases.
AUTHOR: **Stratton C.W.**
CORPORATE SOURCE: Dr. C.W. Stratton, Vanderbilt University School of Med., Nashville, TN 37232, United States

SOURCE: Antimicrobics and Infectious Diseases Newsletter, (1998) 17/2 (9-15).
Refs: 129
ISSN: 1069-417X CODEN: AIDIEX
COUNTRY: United States
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 004 Microbiology
005 General Pathology and Pathological Anatomy
006 Internal Medicine
LANGUAGE: English
SUMMARY LANGUAGE: English
AB In summary, **Chlamydia** species have been shown to infect synovial tissues. The synovial infection produced can be persistent and can cause a cell-mediated inflammatory response directed at **chlamydial** HSPs. Antimicrobial therapies of arthritis syndromes to date have demonstrated some success despite not being optimized for persistent **chlamydial** infection. Clearly, additional work is both needed and warranted to firmly establish the role of **Chlamydia** species in connective tissue diseases. Included in this is the need to optimize antimicrobial therapy against persistent infections caused by **Chlamydia**.

L7 ANSWER 26 OF 30 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 97374392 EMBASE
DOCUMENT NUMBER: 1997374392
TITLE: The importance of **Chlamydia** pneumoniae as a pathogen: The 1996 consensus conference on **Chlamydia** pneumoniae infections.
AUTHOR: File T.M. Jr.; Bartlett J.G.; Cassell G.H.; Gaydos C.A.; Grayston J.T.; Hammerschlag M.R.; Jones R.B.; Kahn J.B.; Marie T.J.; Ramirez J.A.; Saikku P.; Schachter J.; Schumacher H.R.; Stamm W.E.; **Stratton C.W.**; Yu V.L.
CORPORATE SOURCE: T.M. File Jr., Infectious Disease Section, SUMMA Health System, 75 Arch Street, Akron, OH 44304-1430, United States
SOURCE: Infectious Diseases in Clinical Practice, (1997) 6/SUPPL. 2 (S28-S31).
Refs: 26
ISSN: 1056-9103 CODEN: IDCPEY
COUNTRY: United States
DOCUMENT TYPE: Journal; (Short Survey)
FILE SEGMENT: 004 Microbiology
015 Chest Diseases, Thoracic Surgery and Tuberculosis
018 Cardiovascular Diseases and Cardiovascular Surgery
LANGUAGE: English

L7 ANSWER 27 OF 30 BIOTECHDS COPYRIGHT 2002 DERWENT INFO AND ISI
ACCESSION NUMBER: 1996-11724 BIOTECHDS
TITLE: Inducing mucosal immune response by administration of antigen-encoding DNA;
genetic immunization against disease caused by HIV virus, herpes virus, orthomyxo virus, mumps virus, hepatitis virus, Salmonella, Shigella, **Chlamydia** or Helicobacter sp.
AUTHOR: **Mitchell W M**
PATENT ASSIGNEE: Univ.Vanderbilt
LOCATION: Nashville, TN, USA.
PATENT INFO: WO 9621356 18 Jul 1996
APPLICATION INFO: WO 1995-US8374 3 Jul 1995
PRIORITY INFO: US 1995-372429 13 Jan 1995
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: WPI: 1996-341965 [34]
AB A new method for inducing a mucosal immune response in a subject involves

administering to the mucosa of the subject an amount of an antigen-encoding DNA effective to induce a mucosal immune response complexed to a transfection-facilitating lipospermine or lipospermidine. Preferably, administration is nasal, oral, rectal or vaginal and the lipospermine is dioctadecylamidoglycylspermine. The DNA encodes an envelope antigen or an envelope-associated antigen. Also new is a composition of the antigen-encoding DNA complexed to a transfection facilitating lipospermine or lipospermidine. The DNA induces both humoral and cellular responses and allows easy formulation of multiple sequence variants into a single genetic vaccine. The method provides the advantages of live, attenuated vaccines and lacks the threat of reversion to virulence. The antigen-encoding DNA may be cloned into a vector, e.g. plasmid pBR322-based plasmid pHenv. (82pp)

L7 ANSWER 28 OF 30 BIOTECHDS COPYRIGHT 2002 DERWENT INFO AND ISI
ACCESSION NUMBER: 2001-00586 BIOTECHDS
TITLE: Diagnosing and monitoring multiple sclerosis by assaying a test sample for **Chlamydia**, and treating multiple sclerosis by administering anti-**chlamydial** agents such as rifamycins, azalides, macrolides, ketolides; polymerase chain reaction for diagnosis and drug screening method
AUTHOR: **Stratton C W**; **Mitchell W M**; Yao S Y; Bannan J D; Ljunggren-Rose A; Sriram S
PATENT ASSIGNEE: Univ.Vanderbilt
LOCATION: Nashville, TN, USA.
PATENT INFO: WO 2000057187 28 Sep 2000
APPLICATION INFO: WO 2000-US7226 17 Mar 2000
PRIORITY INFO: US 2000-176784 18 Jan 2000; US 1999-125598 19 Mar 1999
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: WPI: 2000-602242 [57]
AB A new method (M1) for diagnosing or monitoring multiple sclerosis in an individual is claimed. It involves assaying a test sample (selected from blood, serum or cerebrospinal fluid) for the presence of **Chlamydia** selected from **Chlamydia** pneumoniae by contacting cultured **Chlamydia**-free indicator cells with the test sample. Also claimed are; a method (M2) for isolating elementary bodies from a receptacle by treating the receptacle with trypsin (EC-3.4.21.4)/EDTA(ethylene diamine tetraacetic acid); a method (M3) for releasing DNA from elementary bodies by incubating them under disulfide reducing conditions and digesting with a protease; a method (M4) for treating an individual with multiple sclerosis; a method (M5) for determining if a candidate compound is a potential drug for the treatment of **Chlamydial** infection-related disease by infecting a non-human animal with **Chlamydia**, administering a candidate compound to the animal, and assaying for the presence of a **chlamydial** infection. The methods are useful for diagnosing and treating multiple sclerosis and for identifying agents which may potentially be useful for treating the disease. (100pp)

L7 ANSWER 29 OF 30 BIOTECHDS COPYRIGHT 2002 DERWENT INFO AND ISI
ACCESSION NUMBER: 2000-10037 BIOTECHDS
TITLE: Isolated nucleic acid for use in diagnostic and analytical methods encodes genomic sequence of **Chlamydia** pneumoniae; method is useful for prophylactic and therapeutic purpose
AUTHOR: Stephens R; **Mitchell W**; Kalman S; Davis R
PATENT ASSIGNEE: Univ.California
LOCATION: Oakland, CA, USA.
PATENT INFO: WO 2000027994 18 May 2000
APPLICATION INFO: WO 1999-US26923 12 Nov 1999
PRIORITY INFO: US 1999-128606 8 Apr 1999; US 1998-108279 12 Nov 1998

DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: WPI: 2000-376516 [32]
AB A new isolated nucleic acid (N1) encoding a **Chlamydia** pneumoniae protein (P1) is claimed. Also claimed are: a probe containing a hybridizing fragment of N1; an isolated nucleic acid acid (N2) that hybridizes under stringent conditions to N1; an expression cassette containing N1 under the transcriptional regulation of a initiation region functional in an expression host, and a transcriptional termination region; a cell containing an expression cassette as part of an extrachromosomal element in the genome of a host cell, and the cellular progeny of the host cell; a method for producing a P1 by growing a cell where the protein is expressed and isolating protein free of other proteins; a purified protein composition containing at least 50 wt.% of P1; and a monoclonal antibody binding to the peptide. The isolated nucleic acid is useful for diagnosing and analytical methods, such as hybridization-based assays or amplification-based assays. The protein may be used for diagnostic purposes, for its enzymatic or structural activity, or as a vaccine. (330pp)

L7 ANSWER 30 OF 30 BIOTECHDS COPYRIGHT 2002 DERWENT INFO AND ISI
ACCESSION NUMBER: 1998-05630 BIOTECHDS
TITLE: Treating **chlamydia** infections using a combination of agents directed against different phases; disease diagnosis by DNA amplification
AUTHOR: **Mitchell W M; Stratton C W**
PATENT ASSIGNEE: Univ.Vanderbilt
LOCATION: Nashville, TN, USA.
PATENT INFO: WO 9806435 19 Feb 1998
APPLICATION INFO: WO 1997-US14402 14 Aug 1997
PRIORITY INFO: US 1996-23921 14 Aug 1996
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: WPI: 1998-159294 [14]
AB A new composition (A) comprises at least 2 antichlamydial agents (I) active against different phases in the **Chlamydia** sp. life cycle. Also new are: a method for identifying agents that inhibit **chlamydial** infection or are active against the cryptic forms of **Chlamydia** sp.; a method fr identifying cells containing these cryptic forms; a method for activating macrophages and monocytes in which the ability to combat infection has been comprised by **chlamydial** infection; a method for detecting **chlamydial** elementary bodies in a sample, which involves treatment with a disulfide reducing agent prior to detecting **chlamydial** DNA by DNA amplification using the polymerase chain reaction; 12 specific peptides derived from various **Chlamydia** spp.; the therapeutic agents identified by the method; therapy of **Chlamydia** sp. infection; and determination of the status of a patient and monitoring treatment for **Chlamydia** sp. infection. (A) is used for therapy of **Chlamydia** pneumoniae disease and related diseases (autoimmune diseases, inflammation, and immunodeficiency conditions. **Chlamydia** spp. have been associated with multiple sclerosis, rheumatoid arthritis, inflammatory bowel disease, etc. (82pp)

FILE 'HOME' ENTERED AT 14:11:26 ON 26 MAR 2002

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:37:19 ; Search time 42.75 Seconds
(without alignments)
26.728 Million cell updates/sec

Title: US-09-709-201-97
Perfect score: 82
Sequence: 1 CFGVKGTTVNANELP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73	89.0	389	2	A43587	major outer membra
2	73	89.0	389	2	D86577	major outer membra
3	69	84.1	389	2	I40739	major outer membra
4	61	74.4	389	2	I40864	major outer membra
5	45	54.9	500	2	B81060	N utilization subs
6	45	54.9	505	2	H81816	N utilisation subs
7	44	53.7	744	2	H71643	cell division prot
8	42	51.2	389	1	MMCWP3	major outer membra
9	42	51.2	455	2	F96817	hypothetical prote
10	42	51.2	708	2	A56163	peptide transport
11	41	50.0	93	1	R5RZ23	ribosomal protein
12	41	50.0	93	1	R5ZM23	ribosomal protein
13	41	50.0	93	1	R5WT23	ribosomal protein
14	41	50.0	604	2	T51936	probable 9-cis-epo
15	41	50.0	605	2	T07123	nine-cis-epoxycaro
16	40	48.8	246	2	D85948	probable actin ygg
17	40	48.8	246	2	A65077	hypothetical 26.6K
18	40	48.8	365	1	G64705	conserved hypothet
19	40	48.8	1235	1	QQBEW4	DNA-binding protei
20	40	48.8	1984	2	A44396	P-type cation tran
21	40	48.8	3343	2	S44887	ZK112.7 protein -
22	39	47.6	215	2	C84098	phosphoglycolate p
23	39	47.6	216	1	G70044	histidine-containi
24	39	47.6	295	2	E83302	probable transcrip
25	39	47.6	335	2	E71520	hypothetical prote
26	39	47.6	640	1	ZYBPT5	tail protein pb5 -
27	39	47.6	773	2	T27382	hypothetical prote
28	38.5	47.0	435	2	S74208	galacturan 1,4-alp
29	38	46.3	168	2	T52480	cytochrome-c oxida
30	38	46.3	224	2	T44745	probable oxidoredu
31	38	46.3	295	2	S67934	branched-chain ket
32	38	46.3	322	2	JQ1191	vanH protein - Ent
33	38	46.3	330	2	C69593	3-methyl-2-oxobuta
34	38	46.3	345	2	G83633	hypothetical prote
35	38	46.3	353	2	C72479	hypothetical prote
36	38	46.3	380	2	G01639	transmembrane prot
37	38	46.3	381	2	T02589	hypothetical prote
38	38	46.3	458	2	D72012	phosphoglucomutase
39	38	46.3	458	2	E86611	phosphoglucomutase
40	38	46.3	473	1	A53036	Ca2+/calmodulin-de
41	38	46.3	529	2	E84813	hypothetical prote
42	38	46.3	776	1	JQ2022	outer layer protei
43	38	46.3	823	2	A96501	hypothetical prote
44	38	46.3	1022	2	T24663	hypothetical prote
45	38	46.3	1063	2	T38732	probable helicase

ALIGNMENTS

RESULT 1
A43587
major outer membrane protein, porin CP0051 precursor [imported] - Chlamydophila pneum
N;Alternate names: MOMP
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 11-May-2000
C;Accession: A43587; A49751; A49216; G72044; F81619
R;Perez Melgosa, M.; Kuo, C.C.; Campbell, L.A.
Infect. Immun. 59, 2195-2199, 1991
A;Title: Sequence analysis of the major outer membrane protein gene of Chlamydia pneu
A;Reference number: A43587; MUID:91244474
A;Accession: A43587
A;Molecule type: DNA
A;Residues: 1-389 <PER>
A;Cross-references: GB:M69230; NID:g144540; PIDN:AAA73071.1; PID:g144541
R;Carter, M.W.; Al-Mahdawi, S.A.H.; Giles, I.G.; Treharne, J.D.; Ward, M.E.; Clarke,
J. Gen. Microbiol. 137, 465-475, 1991
A;Title: Nucleotide sequence and taxonomic value of the major outer membrane protein
A;Reference number: A49751; MUID:91237311
A;Accession: A49751
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <CAR>
A;Cross-references: GB:M64064; GB:M34942; NID:g144534; PIDN:AAA23143.1; PID:g144535
A;Note: isolate IOL-207
R;Gaydos, C.A.; Quinn, T.C.; Bobo, L.D.; Eiden, J.J.
Infect. Immun. 60, 5319-5323, 1992
A;Title: Similarity of Chlamydia pneumoniae strains in the variable domain IV region
A;Reference number: A49216; MUID:93084388
A;Accession: A49216
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 297-352 <GAY>
A;Cross-references: GB:S50607; NID:g260972; PIDN:AAB24363.1; PID:g260973
A;Note: sequence extracted from NCBI backbone (NCBIN:120604, NCBIP:120605)
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: G72044
A;Molecule type: DNA
A;Residues: 1-389 <ARN>
A;Cross-references: GB:AE001652; GB:AE001363; NID:g4376997; PIDN:AAD18834.1; PID:g437
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255
A;Accession: F81619
A;Status: preliminary
A;Molecule type: DNA

```
A;Residues: 1-389 <REA>
A;Cross-references: GB:AE002168; GB:AE002161; NID:g7188982; PIDN:AAF37944.1; PID:g718899
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: ompA; CP0051
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match             89.0%;  Score 73;  DB 2;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 3e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||||||||||||
Db     158 FGVKGTTVNANELP 171


RESULT   2
D86577
major outer membrane protein [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: D86577
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: D86577
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <STO>
A;Cross-references: GB:BA000008; NID:g8979067; PIDN:BAA98902.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: ompA
C;Superfamily: Chlamydia major outer membrane protein


  Query Match             89.0%;  Score 73;  DB 2;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 3e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||||||||||||
Db     158 FGVKGTTVNANELP 171


RESULT   3
I40739
major outer membrane protein precursor - Chlamydophila pneumoniae (strain equine/N16)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
A;Variety: strain equine/N16
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Apr-2000
C;Accession: I40739
R;Storey, C.; Lusher, M.; Yates, P.; Richmond, S.
J. Gen. Microbiol. 139, 2621-2626, 1993
A;Title: Evidence for Chlamydia pneumoniae of non-human origin.
A;Reference number: I40739; MUID:94103736
A;Accession: I40739
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-389 <STO>
A;Cross-references: GB:L04982; NID:g289840; PIDN:AAA17397.1; PID:g289841
C;Comment: On the basis of the major outer membrane protein the authors classified the e
 the sequence of the genome strain CWL029 and strain strain IOL-207. See PIR:A43587.
C;Genetics:
A;Gene: momp
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match             84.1%;  Score 69;  DB 2;  Length 389;
  Best Local Similarity   92.9%;  Pred. No. 0.00016;
  Matches   13;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           |||||||:|||||||
Db     158 FGVKGTSVNANELP 171


RESULT   4
I40864
major outer membrane protein - Chlamydophila psittaci
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
C;Accession: I40864; S33465
R;Girjes, A.A.; Carrick, F.N.; Lavin, M.F.
Gene 138, 139-142, 1994
A;Title: Remarkable sequence relatedness in the DNA encoding the major outer membrane
A;Reference number: I40864; MUID:94171025
A;Accession: I40864
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-389 <RES>
A;Cross-references: EMBL:X72023; NID:g313844; PIDN:CAA50906.1; PID:g313845
C;Superfamily: Chlamydia major outer membrane protein


  Query Match             74.4%;  Score 61;  DB 2;  Length 389;
  Best Local Similarity   85.7%;  Pred. No. 0.0044;
  Matches   12;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||||:| |||||
Db     158 FGVKGTSVAANELP 171


RESULT   5
B81060
N utilization substance protein A NMB1642 [imported] - Neisseria meningitidis (strain
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: B81060
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755
A;Accession: B81060
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-500 <TET>
A;Cross-references: GB:AE002514; GB:AE002098; NID:g7226886; PIDN:AAF41991.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1642
C;Superfamily: Escherichia coli transcription factor nusA; transcription termination


  Query Match             54.9%;  Score 45;  DB 2;  Length 500;
  Best Local Similarity   62.5%;  Pred. No. 4.3;
  Matches   10;  Conservative    2;  Mismatches    2;  Indels    2;  Gaps    1;

Qy       1 CFGVKGTTVNA--NEL 14
           | ||:|: |||  |||
Db     259 CIGVRGSRVNAVSNEL 274


RESULT   6
```

H81816
N utilisation substance protein A NMA1896 [imported] - Neisseria meningitidis (strain Z2
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: H81816
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556
A;Accession: H81816
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-505 <PAR>
A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85117.1; PID:g738053
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: nusA; NMA1896
C;Superfamily: Escherichia coli transcription factor nusA; transcription termination fac

Query Match 54.9%; Score 45; DB 2; Length 505;
Best Local Similarity 62.5%; Pred. No. 4.3;
Matches 10; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

Qy 1 CFGVKGTTVNA--NEL 14
| ||:|: ||| |||
Db 264 CIGVRGSRVNAVSNEL 279

RESULT 7
H71643
cell division protein ftsK homolog (ftsK) RP823 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 02-Mar-2001
C;Accession: H71643
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499
A;Accession: H71643
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-744 <AND>
A;Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15248.1; PID:g386134
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: ftsK; RP823
C;Superfamily: Bacillus subtilis DNA translocase spoIIIE

Query Match 53.7%; Score 44; DB 2; Length 744;
Best Local Similarity 57.1%; Pred. No. 9.9;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 FGVKGTTVNANELP 15
||||| :| |: |
Db 288 FGVKGQIININQGP 301

RESULT 8
MMCWP3
major outer membrane protein precursor - Chlamydophila psittaci (strain S26/3)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 31-Mar-2000
C;Accession: S08770
R;Herring, A.J.; Tan, T.W.; Baxter, S.; Inglis, N.F.; Dunbar, S.
FEMS Microbiol. Lett. 65, 153-158, 1989
A;Title: Sequence analysis of the major outer membrane protein gene of an ovine abortion
A;Reference number: S08770
A;Accession: S08770
A;Molecule type: DNA
A;Residues: 1-389 <HER>
A;Cross-references: EMBL:X51859; NID:g40600; PIDN:CAA36152.1; PID:g40601
C;Superfamily: Chlamydia major outer membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-389/Product: major outer membrane protein #status predicted <MAT>

Query Match 51.2%; Score 42; DB 1; Length 389;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVKGTTVNANELP 15
|||||:: |::||
Db 159 GVKGSSIAADQLP 171

RESULT 9
F96817
hypothetical protein F9K20.12 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: F96817
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: F96817
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-455 <STO>
A;Cross-references: GB:AE005173; NID:g3834312; PIDN:AAC83028.1; GSPDB:GN00141
C;Genetics:
A;Gene: F9K20.12
A;Map position: 1

Query Match 51.2%; Score 42; DB 2; Length 455;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFGVKGTTVN 10
| |||| |||
Db 374 CSGVKGKTVN 383

RESULT 10
A56163
peptide transport protein hPEPT1 - human
C;Species: Homo sapiens (man)
C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 24-Sep-1999
C;Accession: A56163
R;Liang, R.; Fei, Y.J.; Prasad, P.D.; Ramamoorthy, S.; Han, H.; Yang-Feng, T.L.; Hedi
J. Biol. Chem. 270, 6456-6463, 1995
A;Title: Human intestinal H(+)/peptide cotransporter. Cloning, functional expression,
A;Reference number: A56163; MUID:95204429
A;Accession: A56163
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-708 <LIA>
A;Cross-references: GB:U13173; NID:g773587; PIDN:AAB61693.1; PID:g773588
C;Superfamily: peptide transport protein PEPT1

Query Match 51.2%; Score 42; DB 2; Length 708;
Best Local Similarity 46.2%; Pred. No. 22;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy       3 GVKGTTVNANELP 15
           |:|| |::: |:|
Db     524 GIKGFTISSTEIP 536


RESULT  11
R5RZ23
ribosomal protein L23 - rice chloroplast
C;Species: chloroplast Oryza sativa (rice)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 24-Sep-1999
C;Accession: JQ0271; S05151; JA0092
R;Shimada, H.; Whittier, R.F.; Hiratsuka, J.; Maeda, Y.; Hirai, A.; Sugiura, M.
submitted to JIPID, December 1989
A;Reference number: JQ0200
A;Accession: JQ0271
A;Molecule type: DNA
A;Residues: 1-93 <SHI>
A;Experimental source: cv. Nihonbare
R;Hiratsuka, J.; Shimada, H.; Whittier, R.; Ishibashi, T.; Sakamoto, M.; Mori, M.; Kondo
M.
Mol. Gen. Genet. 217, 185-194, 1989
A;Title: The complete sequence of the rice (Oryza sativa) chloroplast genome: intermolec
of the cereals.
A;Reference number: S05080; MUID:89364698
A;Accession: S05151
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-14,'A',16-93 <HIR>
A;Experimental source: cv. Nihonbare
A;Note: this sequence was submitted to EMBL, July 1989
R;Moon, E.; Wu, R.
Gene 70, 1-12, 1988
A;Title: Organization and nucleotide sequence of genes at both junctions between the two
A;Reference number: JA0092; MUID:89196901
A;Accession: JA0092
A;Molecule type: DNA
A;Residues: 1-14,'A',16-45,'X',47-93 <MOO>
A;Cross-references: GB:M22826; NID:g710564; PIDN:AAA84593.1; PID:g710565
A;Note: genes located at the two inverted repeats (IRA and IRB) in the rice chloroplast
l proteins L23 and L2 and S19, lies at the ends of the two IRs near the LSCR
A;Note: the 46th codon is the TAG stop codon in this publication
C;Genetics:
A;Gene: rpl23
A;Genome: chloroplast
C;Superfamily: Escherichia coli ribosomal protein L23
C;Keywords: chloroplast; protein biosynthesis; ribosome


  Query Match             50.0%;  Score 41;  DB 1;  Length 93;
  Best Local Similarity   57.1%;  Pred. No. 3.7;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||   ||:: ||
Db      43 FGVKVVAVNSHRLP 56


RESULT  12
R5ZM23
ribosomal protein L23 - maize chloroplast
C;Species: chloroplast Zea mays (maize)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jun-2000
C;Accession: S01396; S60127; S58596; S58638; S20058
R;McLaughlin, W.E.; Larrinua, I.M.
Nucleic Acids Res. 16, 8183, 1988
A;Title: The sequence of the maize plastid encoded rpl23 locus.
A;Reference number: S01396; MUID:88335565
A;Accession: S01396
A;Molecule type: DNA
A;Residues: 1-93 <MCL>
A;Cross-references: EMBL:X07864; NID:g12417; PIDN:CAA30712.1; PID:g12418
A;Genetics: GEN1
A;Accession: S60127
A;Molecule type: DNA
A;Residues: 1-93 <MCW>
A;Cross-references: EMBL:X07864; NID:g12417; PIDN:CAA30712.1; PID:g12418
A;Genetics: GEN2
R;Maier, R.M.; Neckermann, K.; Igloi, G.L.; Koessel, H.
J. Mol. Biol. 251, 614-628, 1995
A;Title: Complete sequence of the maize chloroplast genome: gene content, hotspots of
A;Reference number: S58531; MUID:95395841
A;Accession: S58596
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-93 <MAI>
A;Cross-references: EMBL:X86563; NID:g902200; PIDN:CAA60330.1; PID:g902265
A;Genetics: GEN2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
A;Accession: S58638
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-93 <MAW>
A;Cross-references: EMBL:X86563; NID:g902200; PIDN:CAA60330.1; PID:g902265
A;Genetics: GEN1
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
R;Hoch, B.; Maier, R.M.; Appel, K.; Igloi, G.L.; Koessel, H.
Nature 353, 178-180, 1991
A;Title: Editing of a chloroplast mRNA by creation of an initiation codon.
A;Reference number: S17874; MUID:91367263
A;Accession: S20058
A;Molecule type: mRNA
A;Residues: 81-93 <HOC>
A;Cross-references: EMBL:X62070; NID:g12463; PIDN:CAA43984.1; PID:g12465
C;Genetics: <GEN1>
A;Gene: rpl23
A;Map position: IR(A); IR(II)
A;Genome: chloroplast
C;Genetics: <GEN2>
A;Gene: rpl23
A;Map position: IR(B); IR(I)
A;Genome: chloroplast
C;Superfamily: Escherichia coli ribosomal protein L23
C;Keywords: chloroplast; protein biosynthesis; ribosome


  Query Match             50.0%;  Score 41;  DB 1;  Length 93;
  Best Local Similarity   57.1%;  Pred. No. 3.7;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||   ||:: ||
Db      43 FGVKVVAVNSHRLP 56


RESULT  13
R5WT23
ribosomal protein L23 - wheat chloroplast
C;Species: chloroplast Triticum aestivum (common wheat)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 22-Jun-1999
C;Accession: S06026
R;Bowman, C.M.; Barker, R.F.; Dyer, T.A.
Curr. Genet. 14, 127-136, 1988
A;Title: In wheat ctDNA, segments of ribosomal protein genes are dispersed repeats, p
A;Reference number: S06025; MUID:89028843
A;Accession: S06026
A;Molecule type: DNA
A;Residues: 1-93 <BOW>
A;Cross-references: EMBL:X12850; NID:g12369; PIDN:CAA31328.1; PID:g12370
C;Genetics:
A;Genome: chloroplast
C;Superfamily: Escherichia coli ribosomal protein L23
C;Keywords: chloroplast; protein biosynthesis; ribosome


  Query Match             50.0%;  Score 41;  DB 1;  Length 93;
  Best Local Similarity   57.1%;  Pred. No. 3.7;
```

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 FGVKGTTVNANELP 15
           ||||   ||:: ||
Db      43 FGVKVVAVNSHRLP 56
```

RESULT 14
T51936
probable 9-cis-epoxycarotenoid dioxygenase [imported] - potato
C;Species: Solanum tuberosum (potato)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T51936
R;Burbidge, A.; Taylor, I.B.; Thompson, A.
submitted to the EMBL Data Library, March 2000
A;Description: Potato putative 9-cis-epoxycarotenoid dioxygenase 1 cDNA.
A;Reference number: Z25874
A;Accession: T51936
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-604 <BUR>
A;Cross-references: EMBL:AJ276244; PIDN:CAB76920.1
C;Genetics:
A;Gene: nced1

Query Match 50.0%; Score 41; DB 2; Length 604;
Best Local Similarity 53.8%; Pred. No. 27;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 FGVKGTTVNANEL 14
           :|  || :|||:|
Db     588 YGFHGTFINANDL 600
```

RESULT 15
T07123
nine-cis-epoxycarotenoid dioxygenase - tomato
N;Alternate names: probable neoxanthin cleavage enzyme
C;Species: Lycopersicon esculentum (tomato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Jun-2000
C;Accession: T07123
R;Burbidge, A.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z15934
A;Accession: T07123
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-605 <BUR>
A;Cross-references: EMBL:Z97215; PIDN:CAB10168.1

Query Match 50.0%; Score 41; DB 2; Length 605;
Best Local Similarity 53.8%; Pred. No. 28;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 FGVKGTTVNANEL 14
           :|  || :|||:|
Db     589 YGFHGTFINANDL 601
```

Search completed: March 26, 2002, 13:37:19
Job time: 53 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:40:13 ; Search time 79.01 Seconds
(without alignments)
27.770 Million cell updates/sec

Title: US-09-709-201-97
Perfect score: 82
Sequence: 1 CFGVKGTTVNANELP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73	89.0	223	2	Q9RB77	Q9rb77 chlamydia p
2	73	89.0	223	2	Q9RB76	Q9rb76 chlamydia p
3	61	74.4	389	2	Q08085	Q08085 chlamydia p
4	47	57.3	388	2	Q9AIK1	Q9aik1 chlamydia p
5	45	54.9	500	2	Q9JYD3	Q9jyd3 neisseria m
6	45	54.9	505	2	Q9JTB6	Q9jtb6 neisseria m
7	44	53.7	744	2	Q9ZCD4	Q9zcd4 rickettsia
8	44	53.7	1059	2	P95633	P95633 rickettsia
9	43	52.4	389	2	Q9APM4	Q9apm4 chlamydophi
10	42	51.2	318	3	Q9HEU6	Q9heu6 emericella
11	42	51.2	341	2	Q9X717	Q9x717 chlamydophi
12	42	51.2	352	2	O69306	O69306 chlamydia p
13	42	51.2	352	2	O69307	O69307 chlamydia p
14	42	51.2	352	2	O70085	O70085 chlamydia p
15	42	51.2	353	2	O69305	O69305 chlamydia p
16	42	51.2	455	10	Q9ZVA2	Q9zva2 arabidopsis
17	42	51.2	605	4	Q9BZ21	Q9bz21 homo sapien
18	42	51.2	708	4	O43641	O43641 homo sapien
19	42	51.2	743	5	Q9VK47	Q9vk47 drosophila
20	41	50.0	254	10	Q9SM99	Q9sm99 oryza sativ
21	41	50.0	323	2	Q9AC82	Q9ac82 staphylococ
22	41	50.0	604	10	Q9M3Z9	Q9m3z9 solanum tub
23	41	50.0	605	10	O24023	O24023 lycopersico
24	40	48.8	365	2	O26021	O26021 helicobacte
25	40	48.8	1024	2	Q9A4J9	Q9a4j9 caulobacter
26	40	48.8	3337	5	Q9TWY4	Q9twy4 caenorhabdi
27	39	47.6	50	2	Q9Z476	Q9z476 bacillus su
28	39	47.6	143	4	Q04490	Q04490 homo sapien
29	39	47.6	215	2	Q9K6Y7	Q9k6y7 bacillus ha
30	39	47.6	216	2	Q9JMQ2	Q9jmq2 bacillus su
31	39	47.6	282	10	Q9FL72	Q9fl72 arabidopsis
32	39	47.6	295	2	Q9I085	Q9i085 pseudomonas
33	39	47.6	335	2	O84396	O84396 chlamydia t
34	39	47.6	381	10	O48576	O48576 arabidopsis
35	39	47.6	605	5	Q9V7L8	Q9v718 drosophila
36	39	47.6	610	5	Q9U6Y9	Q9u6y9 drosophila
37	39	47.6	773	5	Q9XW87	Q9xw87 caenorhabdi
38	39	47.6	2478	5	Q9VXZ5	Q9vxz5 drosophila
39	39	47.6	5233	5	Q9NB71	Q9nb71 drosophila
40	38	46.3	171	2	Q9RDE1	Q9rde1 streptomyce
41	38	46.3	223	10	Q9FMA6	Q9fma6 arabidopsis
42	38	46.3	224	2	O32906	O32906 mycobacteri
43	38	46.3	229	10	Q9SPV5	Q9spv5 nicotiana p
44	38	46.3	273	3	Q9UQY4	Q9uqy4 aspergillus
45	38	46.3	295	2	Q50850	Q50850 myxococcus

ALIGNMENTS

RESULT 1
Q9RB77
ID Q9RB77 PRELIMINARY; PRT; 223 AA.
AC Q9RB77;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MUTANT MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
GN MOMP.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RA Tharp A.C., Mitchell W.M., Stratton C.W., Ding L.-M.;
RT "Presence of viable Chlamydia pneumoniae in fetal calf serum and in
RT epithelial-derived cell lines.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF131230; AAD33511.1; -.
DR InterPro; IPR000604; Chlamydia_OMP.
DR Pfam; PF01308; Chlamydia_OMP; 1.
DR ProDom; PD001717; Chlamydia_OMP; 1.
FT NON_TER 1 1
FT NON_TER 223 223
SQ SEQUENCE 223 AA; 24171 MW; 6D19A6B4D8841496 CRC64;

Query Match 89.0%; Score 73; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FGVKGTTVNANELP 15
||||||||||||||
Db 58 FGVKGTTVNANELP 71

RESULT 2
Q9RB76
ID Q9RB76 PRELIMINARY; PRT; 223 AA.
AC Q9RB76;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MUTANT MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
GN   MOMP.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tharp A.C., Mitchell W.M., Stratton C.W., Ding L.-M.;
RT   "Presence of viable Chlamydia pneumoniae in fetal calf serum and in
RT   epithelial-derived cell lines.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131229; AAD33510.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   VARIANT       6      6       A -> P.
FT   VARIANT      10     10       M -> K.
FT   VARIANT     161    161       V -> A.
FT   NON_TER     223    223
SQ   SEQUENCE   223 AA;  24176 MW;  4B676047B947C00F CRC64;

  Query Match             89.0%;  Score 73;  DB 2;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||||||||||||
Db     158 FGVKGTTVNANELP 171


RESULT   3
Q08085
ID   Q08085      PRELIMINARY;      PRT;   389 AA.
AC   Q08085;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KOALA TYPE 1;
RX   MEDLINE=94171025; PubMed=8125292;
RA   Girjes A.A., Carrick F.N., Lavin M.F.;
RT   "Remarkable sequence relatedness in the DNA encoding the major outer
RT   membrane protein of Chlamydia psittaci (koala type I) and Chlamydia
RT   pneumoniae.";
RL   Gene 138:189-192(1994).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
DR   EMBL; X72023; CAA50906.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   PRINTS; PR01334; CHLAMIDIAOMP.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     23       BY SIMILARITY.
FT   CHAIN        24    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41579 MW;  5DC50E85A6F4E50F CRC64;

  Query Match             74.4%;  Score 61;  DB 2;  Length 389;
  Best Local Similarity   85.7%;  Pred. No. 0.014;
  Matches   12;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||||:| |||||
Db     158 FGVKGTSVAANELP 171


RESULT   4
Q9AIK1
ID   Q9AIK1      PRELIMINARY;      PRT;   388 AA.
AC   Q9AIK1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VS225;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269259; AAK00240.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     19       POTENTIAL.
FT   CHAIN        20    388       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   388 AA;  41573 MW;  8E232D22C9B9948D CRC64;

  Query Match             57.3%;  Score 47;  DB 2;  Length 388;
  Best Local Similarity   69.2%;  Pred. No. 4;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 GVKGTTVNANELP 15
           |||||:| |::||
Db     156 GVKGTSVAADQLP 168


RESULT   5
Q9JYD3
ID   Q9JYD3      PRELIMINARY;      PRT;   500 AA.
AC   Q9JYD3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   N UTILIZATION SUBSTANCE PROTEIN A.
GN   NMB1642.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / SEROGROUP B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002514; AAF41991.1; -.
DR   TIGR; NMB1642; -.
DR   InterPro; IPR000958; KH.
DR   InterPro; IPR003029; S1.
```

```
DR   Pfam; PF00013; KH-domain; 1.
DR   Pfam; PF00575; S1; 1.
DR   SMART; SM00322; KH; 2.
DR   SMART; SM00316; S1; 1.
KW   Complete proteome.
SQ   SEQUENCE   500 AA;  55751 MW;  753FA50DDEF5B774 CRC64;


  Query Match           54.9%;  Score 45;  DB 2;  Length 500;
  Best Local Similarity  62.5%;  Pred. No. 12;
  Matches   10;  Conservative    2;  Mismatches    2;  Indels    2;  Gaps    1;

Qy      1 CFGVKGTTVNA--NEL 14
          | ||:|: |||  |||
Db    259 CIGVRGSRVNAVSNEL 274


RESULT   6
Q9JTB6
ID   Q9JTB6      PRELIMINARY;      PRT;   505 AA.
AC   Q9JTB6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   N UTILISATION SUBSTANCE PROTEIN A.
GN   NUSA OR NMA1896.
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   menigitidis Z2491.";
RL   Nature 404:502-506(2000).
DR   EMBL; AL162757; CAB85117.1; -.
DR   InterPro; IPR000958; KH.
DR   InterPro; IPR003029; S1.
DR   Pfam; PF00013; KH-domain; 1.
DR   Pfam; PF00575; S1; 1.
DR   SMART; SM00322; KH; 2.
DR   SMART; SM00316; S1; 1.
KW   Complete proteome.
SQ   SEQUENCE   505 AA;  56418 MW;  5A0F080DCA99E5D7 CRC64;


  Query Match           54.9%;  Score 45;  DB 2;  Length 505;
  Best Local Similarity  62.5%;  Pred. No. 12;
  Matches   10;  Conservative    2;  Mismatches    2;  Indels    2;  Gaps    1;

Qy      1 CFGVKGTTVNA--NEL 14
          | ||:|: |||  |||
Db    264 CIGVRGSRVNAVSNEL 279


RESULT   7
Q9ZCD4
ID   Q9ZCD4      PRELIMINARY;      PRT;   744 AA.
AC   Q9ZCD4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CELL DIVISION PROTEIN FTSK HOMOLOG (FTSK).
GN   RP823.
OS   Rickettsia prowazekii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=782;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MADRID E;
RX   MEDLINE=99039499; PubMed=9823893;
RA   Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA   Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA   Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT   "The genome sequence of Rickettsia prowazekii and the origin of
RT   mitochondria.";
RL   Nature 396:133-140(1998).
DR   EMBL; AJ235273; CAA15248.1; -.
DR   InterPro; IPR002543; FtsK_SpoIIIE.
DR   Pfam; PF01580; FtsK_SpoIIIE; 1.
KW   Complete proteome.
SQ   SEQUENCE   744 AA;  82819 MW;  C8208B0A9D9E42AC CRC64;


  Query Match           53.7%;  Score 44;  DB 2;  Length 744;
  Best Local Similarity  57.1%;  Pred. No. 27;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 FGVKGTTVNANELP 15
          |||||  :| |: |
Db    288 FGVKGQIIININQGP 301


RESULT   8
P95633
ID   P95633      PRELIMINARY;      PRT;  1059 AA.
AC   P95633;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ROMPA (FRAGMENT).
GN   OMPA.
OS   Rickettsia montana.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=33991;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M5/6;
RA   Raoult D., Fournier P.E., Roux V.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U83447; AAC35183.1; -.
DR   InterPro; IPR003858; rOmpA_rOmpB.
DR   Pfam; PF02708; rOmpA_rOmpB; 1.
FT   NON_TER       1      1
FT   NON_TER    1059   1059
SQ   SEQUENCE   1059 AA;  110213 MW;  DD9EECF128990632 CRC64;


  Query Match           53.7%;  Score 44;  DB 2;  Length 1059;
  Best Local Similarity  61.5%;  Pred. No. 40;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 FGVKGTTVNANEL 14
          ||| ||| || ::
Db    301 FGVAGTTFNATDI 313


RESULT   9
Q9APM4
ID   Q9APM4      PRELIMINARY;      PRT;   389 AA.
AC   Q9APM4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN   OMP1.
OS   Chlamydophila abortus.
```

```
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LLG;
RX   MEDLINE=20569239; PubMed=11119563;
RA   Vretou E., Psarrou E., Kaisar M., Vlisidou I., Salti-Montesanto V.,
RA   Longbottom D.;
RT   "Identification of protective epitopes by sequencing of the major
RT   outer membrane protein gene of a variant strain of Chlamydia psittaci
RT   serotype 1.";
RL   Infect. Immun. 69:607-612(2001).
DR   EMBL; AF272945; AAG53881.1; -.
KW   Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41897 MW;  20513C69C7DBAAF5 CRC64;


  Query Match             52.4%;  Score 43;  DB 2;  Length 389;
  Best Local Similarity   61.5%;  Pred. No. 20;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GVKGTTVNANELP 15
          ||||::| |::||
Db    159 GVKGSSVAADQLP 171


RESULT  10
Q9HEU6
ID   Q9HEU6      PRELIMINARY;      PRT;   318 AA.
AC   Q9HEU6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PTAC BETA (FRAGMENT).
GN   PTAC.
OS   Emericella nidulans (Aspergillus nidulans).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=5072;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89218944; PubMed=2651886;
RA   Arst H.N. Jr., Tollervey D., Caddick M.X.;
RT   "A translocation associated, loss-of-function mutation in the nitrogen
RT   metabolite repression regulatory gene of Aspergillus nidulans can
RT   revert intracistronically.";
RL   Mol. Gen. Genet. 215:364-367(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Conlon H.E., Zadra I., Haas H., Jones M.G., Arst H.N. Jr.,
RA   Caddick M.X.;
RT   "The Aspergillus nidulans GATA transcription factor gene areB encodes
RT   at least three proteins and features three classes of mutation.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF320980; AAG49360.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   318 AA;  34750 MW;  1F2F8776D1C46A3D CRC64;


  Query Match             51.2%;  Score 42;  DB 3;  Length 318;
  Best Local Similarity   46.7%;  Pred. No. 25;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 CFGVKGTTVNANELP 15
          | |: |  || |: |
Db    237 CLGLSGRQVNRNKTP 251


RESULT  11
Q9X717
ID   Q9X717      PRELIMINARY;      PRT;   341 AA.
AC   Q9X717;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydophila abortus.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LW508;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
DR   EMBL; M73040; AAD29103.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER      1      1
FT   NON_TER    341    341
SQ   SEQUENCE   341 AA;  36762 MW;  B5933C9BF6AAF171 CRC64;


  Query Match             51.2%;  Score 42;  DB 2;  Length 341;
  Best Local Similarity   53.8%;  Pred. No. 27;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GVKGTTVNANELP 15
          ||||::: |::||
Db    123 GVKGSSIAADQLP 135


RESULT  12
O69306
ID   O69306      PRELIMINARY;      PRT;   352 AA.
AC   O69306;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 PROTEIN (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM234;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ004874; CAA06183.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER    352    352
SQ   SEQUENCE   352 AA;  37868 MW;  0AEA9B1E099EED41 CRC64;


  Query Match             51.2%;  Score 42;  DB 2;  Length 352;
  Best Local Similarity   53.8%;  Pred. No. 28;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GVKGTTVNANELP 15
          ||||::: |::||
Db    159 GVKGSSIAADQLP 171


RESULT  13
O69307
ID   O69307      PRELIMINARY;      PRT;   352 AA.
```

```
AC   O69307;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 PROTEIN (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM326;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ004875; CAA06184.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  37854 MW;  33589C6D1137CCDB CRC64;


  Query Match             51.2%;  Score 42;  DB 2;  Length 352;
  Best Local Similarity   53.8%;  Pred. No. 28;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 GVKGTTVNANELP 15
           |||||::: |::||
Db     159 GVKGSSIAADQLP 171


RESULT  14
O70085
ID   O70085      PRELIMINARY;      PRT;   352 AA.
AC   O70085;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PMSH1, PM225;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ005618; CAA06625.1; -.
DR   EMBL; AJ005617; CAA06624.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  37854 MW;  391914AD146072CB CRC64;


  Query Match             51.2%;  Score 42;  DB 2;  Length 352;
  Best Local Similarity   53.8%;  Pred. No. 28;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 GVKGTTVNANELP 15
           |||||::: |::||
Db     159 GVKGSSIAADQLP 171


RESULT  15
O69305
ID   O69305      PRELIMINARY;      PRT;   353 AA.
AC   O69305;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 PROTEIN (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OCLH196;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ004873; CAA06182.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER     353    353
SQ   SEQUENCE   353 AA;  37933 MW;  AC7D8FD9FA6E1728 CRC64;


  Query Match             51.2%;  Score 42;  DB 2;  Length 353;
  Best Local Similarity   53.8%;  Pred. No. 28;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 GVKGTTVNANELP 15
           |||||::: |::||
Db     159 GVKGSSIAADQLP 171


Search completed: March 26, 2002, 13:40:14
Job time: 228 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:40:43 ; Search time 24.63 Seconds
(without alignments)
22.329 Million cell updates/sec

Title: US-09-709-201-97
Perfect score: 82
Sequence: 1 CFGVKGTTVNANELP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73	89.0	389	1	OMP1_CHLPN	P27455 chlamydia p
2	69	84.1	389	1	OM1N_CHLPN	Q07430 chlamydia p
3	61	74.4	333	1	OM1K_CHLPN	Q9xbf4 chlamydia p
4	42	51.2	389	1	OM1A_CHLPS	P16567 chlamydia p
5	42	51.2	708	1	PET1_HUMAN	P46059 homo sapien
6	41	50.0	93	1	RK23_MAIZE	P09387 zea mays (m
7	41	50.0	93	1	RK23_ORYSA	P12097 oryza sativ
8	41	50.0	93	1	RK23_WHEAT	P11535 triticum ae
9	40	48.8	246	1	YGGE_ECOLI	P11668 escherichia
10	40	48.8	1235	1	DNBI_HCMVA	P17147 human cytom
11	40	48.8	1956	1	ATX1_PLAFA	Q04956 plasmodium
12	40	48.8	3343	1	YOG7_CAEEL	P34616 caenorhabdi
13	39	47.6	640	1	TPB5_BPT5	P23207 bacteriopha
14	38.5	47.0	435	1	PGLX_ASPTU	Q00293 aspergillus
15	38	46.3	168	1	COX2_THETH	P98052 thermus aqu
16	38	46.3	223	1	GL1C_ARATH	Q9fma6 arabidopsis
17	38	46.3	229	1	NEC1_NICPL	Q9spv5 nicotiana p
18	38	46.3	322	1	VANH_ENTFC	Q05709 enterococcu
19	38	46.3	330	1	ODBA_BACSU	P37940 bacillus su
20	38	46.3	473	1	KCC4_HUMAN	Q16566 homo sapien
21	38	46.3	776	1	VP4_ROTB5	P36305 bovine rota
22	37	45.1	93	1	RK23_ARATH	P27108 arabidopsis
23	37	45.1	93	1	RK23_TOBAC	P06391 nicotiana t
24	37	45.1	232	1	PYRK_ARCFU	O29722 archaeoglob
25	37	45.1	335	1	NAGZ_XYLFA	Q9paz0 xylella fas
26	37	45.1	395	1	NUSA_HELPJ	Q9zja6 helicobacte
27	37	45.1	395	1	NUSA_HELPY	P55977 helicobacte
28	37	45.1	2731	1	RRPB_CVMJH	P29982 murine coro
29	36	43.9	202	1	YJ72_YEAST	P47142 saccharomyc
30	36	43.9	246	1	PSA6_CAEEL	O17586 caenorhabdi
31	36	43.9	413	1	COBL_PSEDE	P21921 pseudomonas
32	36	43.9	425	1	SYH_STREQ	P30053 streptococc
33	36	43.9	474	1	LCND_LACLA	Q00565 lactococcus
34	36	43.9	539	1	VL1_HPV45	P36741 human papil
35	36	43.9	556	1	FTHS_CLOCY	Q07064 clostridium
36	36	43.9	916	1	CAD4_HUMAN	P55283 homo sapien
37	36	43.9	964	1	PMPE_CHLTR	O84877 chlamydia t
38	36	43.9	976	1	PMPE_CHLMU	Q9pl47 chlamydia m
39	36	43.9	1178	1	YNI7_YEAST	P48231 saccharomyc
40	36	43.9	1208	1	RCQ4_HUMAN	O94761 homo sapien
41	36	43.9	1242	1	CY41_TRYBB	Q99279 trypanosoma
42	36	43.9	1271	1	Y338_MYCGE	P47580 mycoplasma
43	36	43.9	2733	1	RRPB_CVMA5	P16342 murine coro
44	35.5	43.3	215	1	HP27_TAMAS	Q06577 tamias asia
45	35.5	43.3	238	1	YO97_CAEEL	P41847 caenorhabdi

ALIGNMENTS

RESULT 1
OMP1_CHLPN
ID OMP1_CHLPN STANDARD; PRT; 389 AA.
AC P27455; Q9JQF6;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN OMPA OR OMP1 OR CPN0695 OR CP0051.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IOL-207;
RX MEDLINE=91237311; PubMed=2033374;
RA Carter M.W., Al-Mahdawi S.A.H., Giles I.G., Treharne J.D.,
RA Ward M.E., Clarke I.N.;
RT "Nucleotide sequence and taxonomic value of the major outer membrane
RT protein gene of Chlamydia pneumoniae IOL-207.";
RL J. Gen. Microbiol. 137:465-475(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TWAR;
RX MEDLINE=91244474; PubMed=1840574;
RA Perez Melgosa M., Kuo C.-C., Campbell L.A.;
RT "Sequence analysis of the major outer membrane protein gene of
RT Chlamydia pneumoniae.";
RL Infect. Immun. 59:2195-2199(1991).
RN [3]
RP SEQUENCE FROM N.A.
RA Mitchell W.M., Tharp A.C., Stratton C.W., Sriram S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;

```
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20298986; PubMed=10839753;
RA   Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,
RA   Takeuchi H., Nishida J., Shibata K., Fujinaga R., Yoneda H.,
RA   Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,
RA   Ishii K., Shiba T., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of outer membrane protein genes omp and pmp in the whole
RT   genome sequences of Chlamydia pneumoniae isolates from Japan and the
RT   United States.";
RL   J. Infect. Dis. 181 Suppl 3:S524-S527(2000).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M64064; AAA23143.1; -.
DR   EMBL; M69230; AAA73071.1; -.
DR   EMBL; AF131889; AAD22492.1; -.
DR   EMBL; AE001652; AAD18834.1; -.
DR   EMBL; AE002167; AAF37944.1; -.
DR   EMBL; AP002547; BAA98902.1; -.
DR   EMBL; AB033787; BAA85940.1; -.
DR   PIR; A43587; A43587.
DR   PIR; A49751; A49751.
DR   TIGR; CP0051; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
FT   SIGNAL        1     23
FT   CHAIN        24    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41620 MW;  15D984151E41F8F2 CRC64;


  Query Match             89.0%;  Score 73;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||||||||||||
Db     158 FGVKGTTVNANELP 171


RESULT   2
OM1N_CHLPN
ID   OM1N_CHLPN     STANDARD;      PRT;   389 AA.
AC   Q07430;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N16;
RX   MEDLINE=94103736; PubMed=8277245;
RA   Storey C., Lusher M., Yates P., Richmond S.;
RT   "Evidence for Chlamydia pneumoniae of non-human origin.";
RL   J. Gen. Microbiol. 139:2621-2626(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L04982; AAA17397.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     23       BY SIMILARITY.
FT   CHAIN        24    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41628 MW;  801622F05D841967 CRC64;


  Query Match             84.1%;  Score 69;  DB 1;  Length 389;
  Best Local Similarity   92.9%;  Pred. No. 7.2e-05;
  Matches   13;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||||:|||||||
Db     158 FGVKGTSVNANELP 171


RESULT   3
OM1K_CHLPN
ID   OM1K_CHLPN     STANDARD;      PRT;   333 AA.
AC   Q9XBF4;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN (MOMP) (FRAGMENT).
GN   OMPA OR OMP1.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KOALA TYPE I;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M73038; AAD38210.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin.
FT   NON_TER       1      1
FT   NON_TER     333    333
SQ   SEQUENCE   333 AA;  35811 MW;  204604512C4C3B3F CRC64;


  Query Match             74.4%;  Score 61;  DB 1;  Length 333;
  Best Local Similarity   85.7%;  Pred. No. 0.0017;
  Matches   12;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 FGVKGTTVNANELP 15
          ||||||:| |||||
Db    114 FGVKGTSVAANELP 127


RESULT   4
OM1A_CHLPS
ID   OM1A_CHLPS     STANDARD;      PRT;   389 AA.
AC   P16567;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OVINE ENZOOTIC ABORTION ISOLATE S26/3;
RX   MEDLINE=90128177; PubMed=2612883;
RA   Herring A.J., Tan T.W., Baxter S., Inglis N.F., Dunbar S.;
RT   "Sequence analysis of the major outer membrane protein gene of an
RT   ovine abortion strain of Chlamydia psittaci.";
RL   FEMS Microbiol. Lett. 53:153-158(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BOVINE ABORTION ISOLATE BA1;
RX   MEDLINE=96189695; PubMed=8605581;
RA   Griffiths P.C., Plater J.M., Martin T.C., Hughes S.L.,
RA   Hughes K.J., Hewinson R.G., Dawson M.;
RT   "Epizootic bovine abortion in a dairy herd: characterization of a
RT   Chlamydia psittaci isolate and antibody response.";
RL   Br. Vet. J. 151:683-693(1995).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X51859; CAA36152.1; -.
DR   EMBL; L39020; AAB02850.1; -.
DR   PIR; S08770; MMCWP3.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41883 MW;  741B5A23ACDBB447 CRC64;


  Query Match             51.2%;  Score 42;  DB 1;  Length 389;
  Best Local Similarity   53.8%;  Pred. No. 5.1;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GVKGTTVNANELP 15
          ||||::: |::||
Db    159 GVKGSSIAADQLP 171


RESULT   5
PET1_HUMAN
ID   PET1_HUMAN     STANDARD;      PRT;   708 AA.
AC   P46059;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   OLIGOPEPTIDE TRANSPORTER, SMALL INTESTINE ISOFORM (PEPTIDE TRANSPORTER
DE   1) (INTESTINAL H+/PEPTIDE COTRANSPORTER) (SOLUTE CARRIER FAMILY 15,
DE   MEMBER 1).
GN   SLC15A1 OR PEPT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Intestine;
RX   MEDLINE=95204429; PubMed=7896779;
RA   Liang R., Fei Y.-J., Prasad P.D., Ramamoorthy S., Han H.,
RA   Yang-Feng T.L., Hediger M.A., Ganapathy V., Leibach F.H.;
RT   "Human intestinal H+/peptide cotransporter. Cloning, functional
RT   expression, and chromosomal localization.";
RL   J. Biol. Chem. 270:6456-6463(1995).
CC   -!- FUNCTION: PROTON-COUPLED INTAKE OF OLIGOPEPTIDES OF 2 TO 4
CC       AMINO ACIDS WITH A PREFERENCE FOR DIPEPTIDES. MAY CONSTITUTE
CC       A MAJOR ROUTE FOR THE ABSORPTION OF PROTEIN DIGESTION END-
CC       PRODUCTS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U13173; AAB61693.1; -.
DR   EMBL; U21936; AAA63797.1; -.
DR   MIM; 600544; -.
DR   InterPro; IPR000109; PTR2.
DR   Pfam; PF00854; PTR2; 2.
DR   PROSITE; PS01022; PTR2_1; 1.
DR   PROSITE; PS01023; PTR2_2; 1.
KW   Peptide transport; Transport; Transmembrane; Symport; Glycoprotein.
FT   TRANSMEM      1     21       POTENTIAL.
FT   DOMAIN       22     53       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     54     74       POTENTIAL.
FT   DOMAIN       75     82       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     83    103       POTENTIAL.
FT   DOMAIN      104    118       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    119    139       POTENTIAL.
FT   DOMAIN      140    161       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM     162    182       POTENTIAL.
FT   DOMAIN       183    198       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     199    219       POTENTIAL.
FT   DOMAIN       220    276       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     277    297       POTENTIAL.
FT   DOMAIN       298    327       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     328    348       POTENTIAL.
FT   DOMAIN       349    361       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     362    382       POTENTIAL.
FT   DOMAIN       383    584       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     585    605       POTENTIAL.
FT   DOMAIN       606    619       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     620    640       POTENTIAL.
FT   DOMAIN       641    645       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     646    666       POTENTIAL.
FT   DOMAIN       667    708       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      50     50       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     404    404       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     408    408       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     439    439       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     509    509       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     514    514       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     562    562       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   708 AA;  78805 MW;  A75475789177A907 CRC64;
```

Query Match 51.2%; Score 42; DB 1; Length 708;
Best Local Similarity 46.2%; Pred. No. 9.7;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy       3 GVKGTTVNANELP 15
           |:|| |::: |:|
Db     524 GIKGFTISSTEIP 536
```

```
RESULT   6
RK23_MAIZE
ID   RK23_MAIZE     STANDARD;      PRT;    93 AA.
AC   P09387;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   CHLOROPLAST 50S RIBOSOMAL PROTEIN L23.
GN   RPL23.
OS   Zea mays (Maize).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88335565; PubMed=3419911;
RA   McLaughlin W.E., Larrinua I.M.;
RT   "The sequence of the maize plastid encoded rpl 23 locus.";
RL   Nucleic Acids Res. 16:8183-8183(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95395841; PubMed=7666415;
RA   Maier R.M., Neckermann K., Igloi G.L., Koessel H.;
RT   "Complete sequence of the maize chloroplast genome: gene content,
RT   hotspots of divergence and fine tuning of genetic information by
RT   transcript editing.";
RL   J. Mol. Biol. 251:614-628(1995).
RN   [3]
RP   SEQUENCE OF 81-93 FROM N.A.
RX   MEDLINE=91367263; PubMed=1653905;
RA   Hoch B., Maier R.M., Appel K., Igloi G.L., Koessel H.;
RT   "Editing of a chloroplast mRNA by creation of an initiation codon.";
RL   Nature 353:178-180(1991).
CC   -!- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X07864; CAA30712.1; -.
DR   EMBL; X86563; CAA60330.1; -.
DR   EMBL; X86563; CAA60370.1; -.
DR   EMBL; X62070; CAA43984.1; -.
DR   PIR; S01396; R5ZM23.
DR   MaizeDB; 66085; -.
DR   Mendel; 4720; ZEAma;rpl23;1.
DR   InterPro; IPR001014; Ribosomal_L23.
DR   Pfam; PF00276; Ribosomal_L23; 1.
DR   ProDom; PD001141; Ribosomal_L23; 1.
DR   PROSITE; PS00050; RIBOSOMAL_L23; 1.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
SQ   SEQUENCE   93 AA;  10737 MW;  720C2D893CF77566 CRC64;
```

Query Match 50.0%; Score 41; DB 1; Length 93;
Best Local Similarity 57.1%; Pred. No. 1.7;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 FGVKGTTVNANELP 15
           ||||   ||:: ||
Db      43 FGVKVVAVNSHRLP 56
```

```
RESULT   7
RK23_ORYSA
ID   RK23_ORYSA     STANDARD;      PRT;    93 AA.
AC   P12097;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   CHLOROPLAST 50S RIBOSOMAL PROTEIN L23.
GN   RPL23.
OS   Oryza sativa (Rice).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Sugiura M.;
RL   Submitted (JUL-1989) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   COMPLETE GENOME.
RX   MEDLINE=89364698; PubMed=2770692;
RA   Hiratsuka J., Shimada H., Whittier R., Ishibashi T., Sakamoto M.,
RA   Mori M., Kondo C., Honji Y., Sun C.-R., Meng B.-Y., Li Y.-Q.,
RA   Kanno A., Nishizawa Y., Hirai A., Shinozaki K., Sugiura M.;
RT   "The complete sequence of the rice (Oryza sativa) chloroplast genome:
RT   intermolecular recombination between distinct tRNA genes accounts for
RT   a major plastid DNA inversion during the evolution of the cereals.";
RL   Mol. Gen. Genet. 217:185-194(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LABELLE; TISSUE=Seedling;
RX   MEDLINE=89196901; PubMed=3240862;
RA   Moon E., Wu R.;
RT   "Organization and nucleotide sequence of genes at both junctions
RT   between the two inverted repeats and the large single-copy region in
RT   the rice chloroplast genome.";
RL   Gene 70:1-12(1988).
CC   -!- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X15901; CAA33938.1; -.
DR   EMBL; X15901; CAA33923.1; -.
DR   EMBL; M22826; AAA84593.1; -.
DR   EMBL; L40578; AAD15253.1; -.
DR   PIR; JQ0271; R5RZ23.
DR   Mendel; 5010; ORYsa;rpl23;2.
DR   InterPro; IPR001014; Ribosomal_L23.
DR   Pfam; PF00276; Ribosomal_L23; 1.
DR   ProDom; PD001141; Ribosomal_L23; 1.
DR   PROSITE; PS00050; RIBOSOMAL_L23; 1.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   CONFLICT     15     15       R -> A (IN REF. 3).
SQ   SEQUENCE   93 AA;  10764 MW;  721EBD893CF77566 CRC64;


  Query Match             50.0%;  Score 41;  DB 1;  Length 93;
  Best Local Similarity   57.1%;  Pred. No. 1.7;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||   ||:: ||
Db      43 FGVKVVAVNSHRLP 56


RESULT   8
RK23_WHEAT
ID   RK23_WHEAT     STANDARD;      PRT;     93 AA.
AC   P11535; P41097;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   CHLOROPLAST 50S RIBOSOMAL PROTEIN L23.
GN   RPL23.
OS   Triticum aestivum (Wheat), and Hordeum vulgare (Barley).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565, 4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=T.aestivum;
RX   MEDLINE=89028843; PubMed=3180271;
RA   Bowman C.M., Barker R.F., Dyer T.A.;
RT   "In wheat ctDNA, segments of ribosomal protein genes are dispersed
RT   repeats, probably conserved by nonreciprocal recombination.";
RL   Curr. Genet. 14:127-136(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=H.vulgare; STRAIN=CV. HAISA;
RX   MEDLINE=95086380; PubMed=7994178;
RA   Hess W.R., Hoch B., Zeltz P., Huebschmann T., Koessel H.,
RA   Boerner T.;
RT   "Inefficient rpl2 splicing in barley mutants with ribosome-deficient
RT   plastids.";
RL   Plant Cell 6:1455-1465(1994).
CC   -!- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X12850; CAA31328.1; -.
DR   EMBL; X78185; CAA55027.1; -.
DR   PIR; S06026; R5WT23.
DR   InterPro; IPR001014; Ribosomal_L23.
DR   Pfam; PF00276; Ribosomal_L23; 1.
DR   ProDom; PD001141; Ribosomal_L23; 1.
DR   PROSITE; PS00050; RIBOSOMAL_L23; 1.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
SQ   SEQUENCE   93 AA;  10746 MW;  1E747D893CF77562 CRC64;


  Query Match             50.0%;  Score 41;  DB 1;  Length 93;
  Best Local Similarity   57.1%;  Pred. No. 1.7;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           ||||   ||:: ||
Db      43 FGVKVVAVNSHRLP 56


RESULT   9
YGGE_ECOLI
ID   YGGE_ECOLI     STANDARD;      PRT;   246 AA.
AC   P11668;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL PROTEIN YGGE.
GN   YGGE OR B2922 OR Z4259 OR ECS3793.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [4]
RP   SEQUENCE OF 1-155 FROM N.A.
RC   STRAIN=K12 / CS520;
RX   MEDLINE=89313302; PubMed=2546007;
RA   Alefounder P.R., Baldwin S.A., Perham S.A., Short N.J.;
RT   "Identification, molecular cloning and sequence analysis of a gene
```

```
RT   cluster encoding the class II fructose 1,6-bisphosphate aldolase, 3-
RT   phosphoglycerate kinase and a putative second glyceraldehyde 3-
RT   phosphate dehydrogenase of Escherichia coli.";
RL   Mol. Microbiol. 3:723-732(1989).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U28377; AAA69089.1; -.
DR   EMBL; AE000375; AAC75959.1; -.
DR   EMBL; AE005522; AAG58048.1; -.
DR   EMBL; AP002563; BAB37216.1; -.
DR   EMBL; X14436; CAA32608.1; -.
DR   PIR; S04737; S04737.
DR   EcoGene; EG11244; yggE.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   246 AA;  26635 MW;  92436CA0B3DFA891 CRC64;


  Query Match             48.8%;  Score 40;  DB 1;  Length 246;
  Best Local Similarity   53.8%;  Pred. No. 7.3;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 GVKGTTVNANELP 15
           |: |    |||||
Db      13 GISGMAAQANELP 25


RESULT 10
DNBI_HCMVA
ID   DNBI_HCMVA     STANDARD;      PRT;  1235 AA.
AC   P17147;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR DNA-BINDING PROTEIN (MDBP).
GN   UL57 OR DBP.
OS   Human cytomegalovirus (strain AD169).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
OX   NCBI_TaxID=10360;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90269039; PubMed=2161319;
RA   Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA   Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA   Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT   "Analysis of the protein-coding content of the sequence of human
RT   cytomegalovirus strain AD169.";
RL   Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC   -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
CC       REPLICATION.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17403; CAA35372.1; -.
DR   PIR; S09820; QQBEW4.
DR   InterPro; IPR000635; Viral_DNA_bind.
DR   Pfam; PF00747; viral_DNA_bp; 1.
KW   DNA-binding; DNA replication; Zinc-finger; Nuclear protein;
KW   Early protein.
FT   ZN_FING     467    481       C2HC-TYPE.
SQ   SEQUENCE   1235 AA;  133878 MW;  94E8D4F8D3BB2CB6 CRC64;


  Query Match             48.8%;  Score 40;  DB 1;  Length 1235;
  Best Local Similarity   52.6%;  Pred. No. 40;
  Matches   10;  Conservative    0;  Mismatches    3;  Indels    6;  Gaps    1;

Qy       1 CFGVKGTT------VNANE 13
           |||| | |      ||| |
Db    1191 CFGVPGETGGGCFLVNAGE 1209


RESULT 11
ATX1_PLAFA
ID   ATX1_PLAFA     STANDARD;      PRT;  1956 AA.
AC   Q04956;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE CATION-TRANSPORTING ATPASE 1 (EC 3.6.3.-).
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=T9/96;
RX   MEDLINE=93132070; PubMed=8421054;
RA   Krishna S., Cowan G., Meade J.C., Wells R.A., Stringer J.R.,
RA   Robson K.J.;
RT   "A family of cation ATPase-like molecules from Plasmodium
RT   falciparum.";
RL   J. Cell Biol. 120:385-398(1993).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2
CC       ATPASES). SUBFAMILY V.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X65738; CAA46646.1; -.
DR   InterPro; IPR001757; E1-E2_ATPase.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
KW   Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.
FT   DOMAIN        1     35       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     36     58       POTENTIAL.
FT   DOMAIN       59     61       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     62     80       POTENTIAL.
FT   DOMAIN       81    407       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    408    427       POTENTIAL.
FT   DOMAIN      428    440       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    441    462       POTENTIAL.
FT   DOMAIN      463   1818       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1819   1837       POTENTIAL.
FT   DOMAIN     1838   1845       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1846   1863       POTENTIAL.
FT   DOMAIN     1864   1881       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1882   1905       POTENTIAL.
FT   DOMAIN     1906   1928       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1929   1952       POTENTIAL.
FT   DOMAIN     1953   1956       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES     496    496       PHOSPHORYLATION (BY SIMILARITY).
FT   METAL      1760   1760       MAGNESIUM (BY SIMILARITY).
FT   METAL      1764   1764       MAGNESIUM (BY SIMILARITY).
```

```
FT   DOMAIN      246    251       POLY-ASN.
FT   DOMAIN      252    256       POLY-LYS.
FT   DOMAIN      937    941       POLY-ASN.
FT   DOMAIN     1344   1347       POLY-LYS.
FT   DOMAIN     1363   1372       POLY-ASN.
FT   DOMAIN     1680   1684       POLY-ASN.
SQ   SEQUENCE   1956 AA;  230285 MW;  AE708AAE99009335 CRC64;


  Query Match            48.8%;  Score 40;  DB 1;  Length 1956;
  Best Local Similarity   63.6%;  Pred. No. 65;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 VKGTTVNANEL 14
          |||| :|:|:|
Db    263 VKGTYINSNDL 273


RESULT  12
YOG7_CAEEL
ID   YOG7_CAEEL     STANDARD;      PRT;  3343 AA.
AC   P34616;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL 375.7 KDA PROTEIN ZK112.7 IN CHROMOSOME III PRECURSOR.
GN   ZK112.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: CONTAINS 11 CADHERIN DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14324; AAA28182.1; -.
DR   PIR; S44887; S44887.
DR   WormPep; ZK112.7; CE00378.
DR   InterPro; IPR002126; Cadherin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001791; Laminin_G.
DR   Pfam; PF00028; cadherin; 11.
DR   Pfam; PF00054; laminin_G; 1.
DR   PRINTS; PR00205; CADHERIN.
DR   SMART; SM00112; CA; 12.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00282; LamG; 1.
DR   PROSITE; PS00232; CADHERIN_1; 8.
DR   PROSITE; PS50268; CADHERIN_2; 11.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW   Hypothetical protein; Cell adhesion; Signal; Transmembrane;
KW   Cytoskeleton; Glycoprotein; Calcium-binding; Repeat.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27   3343       HYPOTHETICAL PROTEIN ZK112.7.
FT   DOMAIN       27   3228       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   3229   3250       POTENTIAL.
FT   DOMAIN     3251   3343       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       28    117       CADHERIN 1.
FT   DOMAIN      118    229       CADHERIN 2.
FT   DOMAIN      242    330       CADHERIN 3.
FT   DOMAIN      632    738       CADHERIN 4.
FT   DOMAIN     1279   1368       CADHERIN 5.
FT   DOMAIN     1545   1648       CADHERIN 6.
FT   DOMAIN     1676   1756       CADHERIN 7.
FT   DOMAIN     1757   1857       CADHERIN 8.
FT   DOMAIN     1954   2045       CADHERIN 9.
FT   DOMAIN     2046   2145       CADHERIN 10.
FT   DOMAIN     2146   2245       CADHERIN 11.
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    149    149       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    250    250       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    288    288       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    369    369       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    467    467       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    612    612       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    752    752       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    806    806       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    941    941       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    966    966       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    970    970       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    985    985       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1042   1042       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1335   1335       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1425   1425       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1429   1429       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1557   1557       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1563   1563       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1597   1597       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1624   1624       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1695   1695       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1702   1702       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1895   1895       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1900   1900       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2053   2053       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2129   2129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2203   2203       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2382   2382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2391   2391       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2410   2410       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2414   2414       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2431   2431       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2527   2527       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2530   2530       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2564   2564       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2621   2621       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2665   2665       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2712   2712       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2798   2798       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2809   2809       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2927   2927       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2976   2976       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3045   3045       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3222   3222       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3225   3225       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   3343 AA;  375745 MW;  063E6F17FCC15D18 CRC64;


  Query Match            48.8%;  Score 40;  DB 1;  Length 3343;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy       2 FGVKGTTVNANELP 15
           ||::  |:| | ||
Db     229 FGIRSLTLNWNGLP 242


RESULT  13
TPB5_BPT5
ID   TPB5_BPT5      STANDARD;      PRT;   640 AA.
AC   P23207;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   TAIL PROTEIN PB5.
GN   OAD.
OS   Bacteriophage T5.
OC   Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
OX   NCBI_TaxID=10726;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91123205; PubMed=1825083;
RA   Krauel V., Heller K.J.;
RT   "Cloning, sequencing, and recombinational analysis with bacteriophage
RT   BF23 of the bacteriophage T5 oad gene encoding the receptor-binding
RT   protein.";
RL   J. Bacteriol. 173:1287-1297(1991).
RN   [2]
RP   SEQUENCE OF 1-21 FROM N.A.
RX   MEDLINE=94335651; PubMed=8057856;
RA   Decker K., Krauel V., Meesmann A., Heller K.J.;
RT   "Lytic conversion of Escherichia coli by bacteriophage T5: blocking
RT   of the FhuA receptor protein by a lipoprotein expressed early during
RT   infection.";
RL   Mol. Microbiol. 12:321-332(1994).
CC   -!- FUNCTION: STRUCTURAL COMPONENT OF THE TAIL. MEDIATES T5 BINDING TO
CC       ITS HOST RECEPTOR, THE E.COLI FHUA PROTEIN.
CC   -!- MISCELLANEOUS: ONLY SMALL REGIONS OF PB5 MAY BE INVOLVED IN
CC       PROTEIN-PROTEIN INTERACTIONS IN THE TAIL STRUCTURE, THE MAJOR PART
CC       BEING USED IN RECEPTOR BINDING.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M62847; AAA32559.1; -.
DR   EMBL; X75922; CAA53526.1; -.
DR   PIR; A38181; ZYBPT5.
KW   Structural protein; Late protein.
SQ   SEQUENCE   640 AA;  68756 MW;  7EE77F287D2C1DE3 CRC64;


  Query Match             47.6%;  Score 39;  DB 1;  Length 640;
  Best Local Similarity   61.5%;  Pred. No. 30;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANEL 14
           | :||| ||| :|
Db     560 FTIKGTVVNACKL 572


RESULT  14
PGLX_ASPTU
ID   PGLX_ASPTU     STANDARD;      PRT;   435 AA.
AC   Q00293;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   EXOPOLYGALACTURONASE PRECURSOR (EC 3.2.1.67) (EXOPG) (GALACTURAN 1,4-
DE   ALPHA-GALACTURONIDASE) (POLY(1,4-ALPHA-D-
DE   GALACTURONIDE)GALACTURONOHYDROLASE).
GN   PGAX.
OS   Aspergillus tubingensis.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=5068;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 23-41; 172-191 AND 301-315.
RX   MEDLINE=97008968; PubMed=8856078;
RA   Kester H.C.M., Kusters-Van Someren M.A., Mueller Y., Visser J.;
RT   "Primary structure and characterization of an exopolygalacturonase
RT   from Aspergillus tubingensis.";
RL   Eur. J. Biochem. 240:738-746(1996).
CC   -!- FUNCTION: THE ACTIVITY OF THIS ENZYME IS OPTIMAL AT PH 4.2.
CC       SPECIFIC IN HYDROLYZING THE TERMINAL GLYCOSIDIC BOND OF
CC       POLYGALACTURONIC ACID AND OLIGOGALACTURONATES.
CC   -!- CATALYTIC ACTIVITY: RANDOM HYDROLYSIS OF 1,4-ALPHA-D-
CC       GALACTOSIDURONIC LINKAGES IN PECTATE AND OTHER GALACTURONANS.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
CC       (POLYGALACTURONASES).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X99795; CAA68128.1; -.
DR   InterPro; IPR000743; Polygalacturonase.
DR   Pfam; PF00295; Glyco_hydro_28; 1.
DR   PROSITE; PS00502; POLYGALACTURONASE; 1.
KW   Hydrolase; Glycosidase; Cell wall; Signal; Glycoprotein.
FT   SIGNAL        1     22
FT   CHAIN        23    435       EXOPOLYGALACTURONASE.
FT   ACT_SITE    267    267       PROBABLE.
FT   CARBOHYD     93     93       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    128    128       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    198    198       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    252    252       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    264    264       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    296    296       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    328    328       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    353    353       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    406    406       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    429    429       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     25     25       R -> L (IN REF. 1; AA SEQUENCE).
FT   CONFLICT     31     31       C -> T (IN REF. 1; AA SEQUENCE).
FT   CONFLICT     41     41       P -> L (IN REF. 1; AA SEQUENCE).
SQ   SEQUENCE   435 AA;  47296 MW;  C47CD97FCD1657C0 CRC64;


  Query Match             47.0%;  Score 38.5;  DB 1;  Length 435;
  Best Local Similarity   53.3%;  Pred. No. 25;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       1 CFGVKGTTVNANELP 15
           |:| | ||:  || |
Db     348 CYGQKNTTL-CNEYP 361


RESULT  15
COX2_THETH
ID   COX2_THETH     STANDARD;      PRT;   168 AA.
AC   P98052;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE II (EC 1.9.3.1) (CYTOCHROME C BA(3)
```

```
DE   SUBUNIT II) (CYTOCHROME CBA3 SUBUNIT 2).
GN   CBAB OR CTAC.
OS   Thermus aquaticus (subsp. thermophilus).
OC   Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX   NCBI_TaxID=274;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-30.
RC   STRAIN=HB8 / ATCC 27634;
RX   MEDLINE=95386472; PubMed=7657607;
RA   Keightley J.A., Zimmermann B.H., Mather M.W., Springer P.,
RA   Pastuszyn A., Lawrence D.M., Fee J.A.;
RT   "Molecular genetic and protein chemical characterization of the
RT   cytochrome ba3 from Thermus thermophilus HB8.";
RL   J. Biol. Chem. 270:20345-20358(1995).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX   MEDLINE=99287095; PubMed=10360350;
RA   Williams P.A., Blackburn N.J., Sanders D., Bellamy H., Stura E.A.,
RA   Fee J.A., McRee D.E.;
RT   "The CuA domain of Thermus thermophilus ba3-type cytochrome c oxidase
RT   at 1.6-A resolution.";
RL   Nat. Struct. Biol. 6:509-516(1999).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RC   STRAIN=HB8 / ATCC 27634;
RX   MEDLINE=20237613; PubMed=10775261;
RA   Soulimane T., Buse G., Bourenkov G.P., Bartunik H.D., Huber R.,
RA   Than M.E.;
RT   "Structure and mechanism of the aberrant ba3-cytochrome c oxidase
RT   from Thermus thermophilus.";
RL   EMBO J. 19:1766-1776(2000).
CC   -!- FUNCTION: SUBUNIT I AND II FORM THE FUNCTIONAL CORE OF THE ENZYME
CC       COMPLEX. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA
CC       HEME A AND CU(A) TO THE BINUCLEAR CENTER FORMED BY HEME A3 AND
CC       CU(B).
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L09121; AAB00369.1; -.
DR   PDB; 2CUA; 28-MAY-99.
DR   InterPro; IPR001505; COX2.
DR   ProDom; PD000131; COX2; 1.
DR   PROSITE; PS00078; COX2; 1.
KW   Oxidoreductase; Respiratory chain; Electron transport; Transmembrane;
KW   Copper; 3D-structure.
FT   DOMAIN        1      3       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM      4     38
FT   DOMAIN       39     69       CYTOPLASMIC (POTENTIAL).
FT   METAL       114    114       COPPER A.
FT   METAL       149    149       COPPER A.
FT   METAL       153    153       COPPER A.
FT   METAL       157    157       COPPER A.
SQ   SEQUENCE   168 AA;  18563 MW;  FE5689FB7672CF05 CRC64;


  Query Match             46.3%;  Score 38;  DB 1;  Length 168;
  Best Local Similarity   50.0%;  Pred. No. 11;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNANELP 15
           | |:|| :|   ||
Db     116 FHVEGTNINVEVLP 129
```

Search completed: March 26, 2002, 13:40:44
Job time: 258 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:38:46 ; Search time 81.51 Seconds
(without alignments)
13.631 Million cell updates/sec

Title: US-09-709-201-97
Perfect score: 82
Sequence: 1 CFGVKGTTVNANELP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	82	100.0	15	20	AAW95324	Peptide fragment o
2	73	89.0	100	20	AAW95295	Chlamydial major o
3	73	89.0	343	20	AAY56771	C. trachomatis ser
4	67	81.7	391	20	AAY35319	Chlamydia pneumoni
5	61	74.4	20	20	AAW84463	Peptide Cp2A deriv
6	61	74.4	20	20	AAW84557	Peptide C.p.2A der
7	42	51.2	257	21	AAG30954	Arabidopsis thalia
8	42	51.2	365	21	AAG30953	Arabidopsis thalia
9	42	51.2	389	20	AAW98188	Chlamydia psittaci
10	42	51.2	455	21	AAG30952	Arabidopsis thalia
11	42	51.2	708	20	AAW74088	Human hPEPT1 prote
12	42	51.2	708	20	AAW74087	Gastro-intestinal
13	42	51.2	708	20	AAW83394	Human protein-coup
14	41	50.0	605	22	AAE04789	Lycopersicon escul
15	41	50.0	605	22	AAB72308	Neoxanthin cleavag
16	40	48.8	246	21	AAB15923	E. coli proliferat
17	39	47.6	352	20	AAY36781	Chlamydia trachoma
18	38.5	47.0	452	15	AAR59792	Aspergillus tubige
19	38	46.3	15	20	AAW95325	Peptide fragment o
20	38	46.3	146	22	AAM18041	Peptide #4475 enco
21	38	46.3	146	22	AAM30554	Peptide #4591 enco
22	38	46.3	146	22	AAM05680	Peptide #4362 enco
23	38	46.3	322	13	AAR24297	Glycopeptide resis
24	38	46.3	380	18	AAW09406	Transforming growt
25	38	46.3	395	20	AAY35621	Chlamydia pneumoni
26	38	46.3	989	20	AAY37242	Chlamydia trachoma
27	38	46.3	2408	13	AAR24307	Translation of ORF
28	37	45.1	89	18	AAW20662	H. pylori secreted
29	37	45.1	97	21	AAB16756	Bacteriophage Dp-1
30	37	45.1	98	18	AAW20457	H. pylori secreted
31	37	45.1	158	22	AAG76134	Human colon cancer
32	37	45.1	395	19	AAW98760	H. pylori GHPO 108
33	37	45.1	463	22	AAG90216	C glutamicum prote
34	37	45.1	469	12	AAR15510	Tomato ACC synthas
35	37	45.1	469	22	AAE00987	Tomato 1-aminocycl
36	37	45.1	469	22	AAB59726	Tomato ACC synthas
37	37	45.1	583	22	AAE04782	Arabidopsis thalia
38	37	45.1	1073	21	AAB01837	Haemophilus influe
39	37	45.1	1079	21	AAB01836	Haemophilus influe
40	36.5	44.5	537	22	AAE05684	Arabidopsis thalia
41	36	43.9	54	13	AAR20077	Sequence encoded b
42	36	43.9	144	20	AAY08240	Human cadherin-4 p
43	36	43.9	144	21	AAG40800	Zea mays protein f
44	36	43.9	147	21	AAG49588	Arabidopsis thalia
45	36	43.9	158	21	AAG49587	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW95324
ID AAW95324 standard; Protein; 15 AA.
XX
AC AAW95324;
XX
DT 15-MAR-1999 (first entry)
XX
DE Peptide fragment of C. pneumoniae CPN158-171.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia pneumoniae.
XX
PN WO9850074-A2.
XX
PD 12-NOV-1998.
XX
PF 06-MAY-1998; 98WO-US09237.
XX
PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
PR 18-FEB-1998; 98US-0025176.
XX

```
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Mitchell WM,  Stratton CW;
XX
DR   WPI; 1999-059653/05.
XX
PT   Composition with two agents effective against different stages of
PT   chlamydial life cycle - comprises agent targetted against cryptic
PT   phase, against elementary body phase, against replicating phase,
PT   probenicid and antiporphyric
XX
PS   Claim 4; Fig 4; 138pp; English.
XX
CC   The invention relates to the diagnosis and management of infections by
CC   Chlamydia species. The invention provides a composition that comprises
CC   at least two agents, where each of the agents is effective against a
CC   different phase of the chlamydial life cycle. The agents are selected
CC   from: (a) agents targetted against cryptic phase of chlamydial life
CC   cycle; (b) agents targetted against elementary body phase of chlamydial
CC   life cycle; (c) agents targetted against replicating phase of chlamydial
CC   life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC   is used to elicit a protective immune response to Chlamydia infection in
CC   an animal or human and is applied until the animal or human tests
CC   negative for Chlamydia infection. It is also used to treat biological
CC   material infected with Chlamydia. Diagnostic kits for antibody assays
CC   against recombinant major outer membrane protein (MOMP), and for DNA
CC   amplification assays for chlamydial genes, are used to diagnose disease,
CC   e.g. autoimmune disease, an inflammatory disease or a disease that
CC   occurs in an immuno-compromised individual, associated with Chlamydia
CC   infection. The kits are used to detect chlamydial elementary bodies in a
CC   sample. They are also used to monitor and/or modify the course of therapy
CC   in a patient. The treatment reduces the acellular load of infectious
CC   Ebstein Barr virus. The method is also used to treat porphyria, by
CC   reducing the number of elementary bodies and applying a drug, e.g.
CC   cimetidine, and antioxidants, to reduce the adverse effects associated
CC   with porphyria. Sequences AAW95324 to AAW95327 represent peptides
CC   employed for the construction of peptide based ELISAs with species
CC   specificity for variable domain 1 (VD1).
XX
SQ   Sequence   15 AA;


  Query Match             100.0%;  Score 82;  DB 20;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-08;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CFGVKGTTVNANELP 15
          |||||||||||||||
Db      1 cfgvkgttvnanelp 15


RESULT   2
AAW95295
ID   AAW95295 standard; Protein; 100 AA.
XX
AC   AAW95295;
XX
DT   15-MAR-1999  (first entry)
XX
DE   Chlamydial major outer membrane protein (MOMP) PN fragment.
XX
KW   Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW   antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW   major outer membrane protein; autoimmune; inflammatory; porphyria;
KW   Ebstein Barr virus; antioxidant.
XX
OS   Chlamydia sp.
XX
PN   WO9850074-A2.
XX
PD   12-NOV-1998.
XX
PF   06-MAY-1998;   98WO-US09237.
XX
PR   18-FEB-1998;   98US-0025521.
PR   06-MAY-1997;   97US-0045689.
PR   06-MAY-1997;   97US-0045739.
PR   06-MAY-1997;   97US-0045779.
PR   06-MAY-1997;   97US-0045780.
PR   06-MAY-1997;   97US-0045784.
PR   06-MAY-1997;   97US-0045787.
PR   14-AUG-1997;   97US-0911593.
PR   18-FEB-1998;   98US-0025174.
PR   18-FEB-1998;   98US-0025176.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Mitchell WM,  Stratton CW;
XX
DR   WPI; 1999-059653/05.
XX
PT   Composition with two agents effective against different stages of
PT   chlamydial life cycle - comprises agent targetted against cryptic
PT   phase, against elementary body phase, against replicating phase,
PT   probenicid and antiporphyric
XX
PS   Disclosure; Fig 1A; 138pp; English.
XX
CC   The invention relates to the diagnosis and management of infections by
CC   Chlamydia species. The invention provides a composition that comprises
CC   at least two agents, where each of the agents is effective against a
CC   different phase of the chlamydial life cycle. The agents are selected
CC   from: (a) agents targetted against cryptic phase of chlamydial life
CC   cycle; (b) agents targetted against elementary body phase of chlamydial
CC   life cycle; (c) agents targetted against replicating phase of chlamydial
CC   life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC   is used to elicit a protective immune response to Chlamydia infection in
CC   an animal or human and is applied until the animal or human tests
CC   negative for Chlamydia infection. It is also used to treat biological
CC   material infected with Chlamydia. Diagnostic kits for antibody assays
CC   against recombinant major outer membrane protein (MOMP), and for DNA
CC   amplification assays for chlamydial genes, are used to diagnose disease,
CC   e.g. autoimmune disease, an inflammatory disease or a disease that
CC   occurs in an immuno-compromised individual, associated with Chlamydia
CC   infection. The kits are used to detect chlamydial elementary bodies in a
CC   sample. They are also used to monitor and/or modify the course of therapy
CC   in a patient. The treatment reduces the acellular load of infectious
CC   Ebstein Barr virus. The method is also used to treat porphyria, by
CC   reducing the number of elementary bodies and applying a drug, e.g.
CC   cimetidine, and antioxidants, to reduce the adverse effects associated
CC   with porphyria. Sequences AAW95272 to AAW95319 represent peptide
CC   fragments of various Chlamydial MOMPs.
XX
SQ   Sequence   100 AA;


  Query Match             89.0%;  Score 73;  DB 20;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 FGVKGTTVNANELP 15
          ||||||||||||||
Db     62 fgvkgttvnanelp 75


RESULT   3
AAY56771
ID   AAY56771 standard; Protein; 343 AA.
XX
AC   AAY56771;
XX
DT   22-FEB-2000  (first entry)
XX
DE   C. trachomatis serovar HuPn MOMP sequence.
XX
KW   Major outer membrane protein; MOMP; Chlamydia; vaccine; immune response;
```

KW cellular response; immunogen; Th1-like CD4 response; mucosal immunity.
XX
OS Chlamydia trachomatis.
XX
PN WO9951745-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-CA00292.
XX
PR 07-APR-1998; 98US-0055765.
XX
PA (UYMA-) UNIV MANITOBA.
XX
PI Bruhnam RC;
XX
DR WPI; 1999-620205/53.
XX
PT Non-replicating vector encoding fragments of the outer membrane protein
PT of Chlamydia, useful in vaccines and as immunogen -
XX
PS Disclosure; Fig 10 A-F; 52pp; English.
XX
CC The invention provides a non-replicating vector that comprises, linked
CC to a promoter, a nucleotide sequence that encodes a region containing at
CC least one of the conserved domains 2, 3 and 5 of a major outer membrane
CC protein (MOMP) of a Chlamydia strain. The vector is used: (a) in
CC vaccines to generate a protective immune response (mainly cellular)
CC against MOMP, and (b) as immunogens to raise anti-MOMP antibodies, useful
CC in standard immunoassays. Immunization with the vector induces a broad
CC spectrum of immune responses, including Th1-like CD4 responses and
CC mucosal immunity, providing significant protection against subsequent
CC challenge. Sequences AAY56757-71 represent MOMP sequences from a variety
CC of serovars of C. trachomatis.
XX
SQ Sequence 343 AA;

Query Match 89.0%; Score 73; DB 20; Length 343;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FGVKGTTVNANELP 15
||||||||||||||
Db 158 fgvkgttvnanelp 171

RESULT 4
AAY35319
ID AAY35319 standard; Protein; 391 AA.
XX
AC AAY35319;
XX
DT 13-SEP-1999 (first entry)
XX
DE Chlamydia pneumoniae transmembrane protein sequence.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-IB01890.
XX
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
PA (GEST) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae
XX
PS Page 1130-1131; Disclosure; 1912pp; English.
XX
CC AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotides sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 391 AA;

Query Match 81.7%; Score 67; DB 20; Length 391;
Best Local Similarity 92.9%; Pred. No. 0.00058;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FGVKGTTVNANELP 15
|||||||||| ||
Db 160 fgvkgttvnanxlp 173

RESULT 5
AAW84463
ID AAW84463 standard; peptide; 20 AA.
XX
AC AAW84463;
XX
DT 23-MAR-1999 (first entry)
XX
DE Peptide Cp2A derived from a major outer membrane protein.
XX
KW Variable domain; major outer membrane protein; MOMP;
KW Chlamydia; detection; infection; vaccine.
XX
OS Synthetic.
OS Chlamydia pneumoniae.
XX
PN WO9857981-A2.
XX
PD 23-DEC-1998.
XX
PF 15-JUN-1998; 98WO-IL00277.
XX
PR 19-JUN-1997; 97IL-0121114.
XX
PA (SAVY-) SAVYON DIAGNOSTICS LTD.
XX
PI Ohana B;
XX
DR WPI; 1999-080945/07.
XX
PT New peptides derived from Chlamydia pneumoniae MOMP protein - useful
PT to detect C. pneumoniae infection
XX
PS Claim 2; Page 53; 39pp; English.
XX
CC The present peptide is derived from the variable domain of the
CC major outer membrane protein (MOMP) of Chlamydia pneumoniae. The
CC peptide is able to react with antibodies formed during C. pneumoniae
CC infection and characterised by having essentially very low
CC cross-reactivity towards antibodies against other Chlamydia species.

```
CC   A mixture of such peptides (see also AAW84462-68) is used to detect
CC   C. pneumoniae infection, and in the preparation of vaccines.
XX
SQ   Sequence   20 AA;


  Query Match             74.4%;  Score 61;  DB 20;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.00024;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 VKGTTVNANELP 15
          ||||||||||||
Db      1 vkgttvnanelp 12


RESULT   6
AAW84557
ID   AAW84557 standard; peptide; 20 AA.
XX
AC   AAW84557;
XX
DT   26-MAR-1999  (first entry)
XX
DE   Peptide C.p.2A derived from the major outer membrane protein.
XX
KW   Variable domain; immunodominant; major outer membrane protein; MOMP;
KW   anti-MOMP antibody; Chlamydia; vaccine; C. trachomatis.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9900414-A1.
XX
PD   07-JAN-1999.
XX
PF   15-JUN-1998;   98WO-IL00276.
XX
PR   19-JUN-1997;   97IL-0121115.
XX
PA   (SAVY-) SAVYON DIAGNOSTICS LTD.
XX
PI   Ohana B;
XX
DR   WPI; 1999-095677/08.
XX
PT   Chlamydia trachomatis specific peptides useful in diagnostic assays
PT   - derived from major outer membrane protein variable domains and
PT   useful in mixtures to detect infection with or immunise against all
PT   serovars
XX
PS   Example 1; Page 24; 78pp; English.
XX
CC   The present sequence represents a peptide derived from variable
CC   domain 2 (VDII) of the Chlamydia pneumoniae major outer membrane
CC   protein (MOMP). The specification also describes C. trachomatis
CC   MOMP derived peptides which have specificity only to C. trachomatis
CC   anti-MOMP antibodies and are non-cross reactive with anti-MOMP
CC   antibodies of other Chlamydia species. Such peptides are useful to
CC   detect C. trachomatis infections in humans. Mixtures of MOMP peptide
CC   mixtures allow detection of and vaccination against all C. trachomatis
CC   serovars, which is not possible with existing MOMP-derived peptides
CC   for C. trachomatis-specific detection.
XX
SQ   Sequence   20 AA;


  Query Match             74.4%;  Score 61;  DB 20;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.00024;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 VKGTTVNANELP 15
          ||||||||||||
Db      1 vkgttvnanelp 12


RESULT   7
AAG30954
ID   AAG30954 standard; Protein; 257 AA.
XX
AC   AAG30954;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 37095.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
```

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 51.2%; Score 42; DB 21; Length 257;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 CFGVKGTTVN 10
           | |||| |||
Db     176 csgvkgktvn 185
```

RESULT 8
AAG30953
ID AAG30953 standard; Protein; 365 AA.
XX
AC AAG30953;
XX

DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37094.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 51.2%; Score 42; DB 21; Length 365;
Best Local Similarity 80.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 CFGVKGTTVN 10
           | |||| |||
Db     284 csgvkgktvn 293
```

RESULT 9
AAW98188
ID AAW98188 standard; Protein; 389 AA.
XX
AC AAW98188;
XX
DT 05-JUL-1999 (first entry)
XX
DE Chlamydia psittaci major outer membrane protein.
XX
KW Major outer membrane protein; MOMP; psittacosis; infection;
KW vaccine; genetic immunisation.
XX
OS Chlamydia psittaci.
XX
PN WO9910005-A1.
XX
PD 04-MAR-1999.
XX
PF 28-AUG-1998; 98WO-US17943.
XX
PR 28-AUG-1997; 97US-0057147.
XX
PA (LOUU) UNIV LOUISIANA & AGRIC & MECH COLLEGE.
XX
PI Baghian A, Chouljenko VN, Kousoulas KG, Tully TN;
XX
DR WPI; 1999-254214/21.
DR N-PSDB; AAX25047.
XX
PT A new vaccine for Chlamydia psittaci infections
XX
PS Disclosure; Page 60-61; 72pp; English.
XX
CC The present sequence is the major outer membrane protein (MOMP)
CC of Chlamydia psittaci strain B577. A claimed MOMP polypeptide (see
CC AAW98184) comprises regions VD3 and VD4 of B577 MOMP, i.e. it lacks
CC regions VD1 and VD2. A claimed vaccine composition includes MOMP
CC polypeptide lacking VD1 and VD2, optionally fused to a maltose
CC binding protein. Also claimed are an isolated nucleic acid
CC encoding the polypeptide, a vector, and a method of preventing C.
CC psittaci infection by administering the vaccine containing the
CC MOMP polypeptide. Vectors encoding MOMP polypeptides lacking
CC regions VD1 and VD2 are useful for genetic vaccination. The
CC vaccines are used to prevent C. psittaci infection, especially in
CC birds.
XX
SQ Sequence 389 AA;

Query Match 51.2%; Score 42; DB 20; Length 389;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 GVKGTTVNANELP 15
           |||||::: |::||
Db     159 gvkgssiaadqlp 171
```

RESULT 10
AAG30952
ID AAG30952 standard; Protein; 455 AA.
XX
AC AAG30952;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37093.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 51.2%; Score 42; DB 21; Length 455;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFGVKGTTVN 10
| |||| |||
Db 374 csgvkgktvn 383

RESULT 11
AAW74088
ID AAW74088 standard; Protein; 708 AA.
XX
AC AAW74088;
XX
DT 04-MAY-1999 (first entry)
XX
DE Human hPEPT1 protein.
XX
KW Gastro-intestinal transport receptor; binding protein; hSI; HPT1;
KW D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
KW intestinal peptide-associated transporter; hypertension; diabetes;
KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KW therapeutic agent delivery; therapy; probe.
XX
OS Homo sapiens.
XX
PN WO9851325-A2.
XX
PD 19-NOV-1998.
XX
PF 15-MAY-1998; 98WO-US10088.
XX
PR 15-MAY-1997; 97US-0046595.
XX
PA (CYTO-) CYTOGEN CORP.
PA (ELAN-) ELAN CORP PLC.
XX
PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
PI Omahony DJ, Patterson CA, Singleton J;
XX
DR WPI; 1999-009568/01.
XX
PT New proteins that bind specifically to receptors in the
PT gastro-intestinal tract and related nucleic acid - chimaeras and
PT antibodies, used to deliver therapeutic or diagnostic agents to, or
PT through, the gastrointestinal tract, e.g. insulin or leuprolide
XX
PS Disclosure; Fig 1; 294pp; English.
XX
CC This sequence is the human hPEPT1 protein. The invention relates to
CC purified proteins (I) that bind specifically to at least one of the
CC gastro-intestinal (GI) tract receptors human intestinal
CC peptide-associated transporter (HPT1), hPEPT1, D2H and human
CC sucrose-isomaltase complex (hSI). (I) provide active transport of
CC therapeutic agents through human and animal GI tissue (into the blood)
CC for in vivo delivery, particularly for treatment or prevention
CC of hypertension, diabetes, osteoporosis, haemophilia, anaemia, cancer,
CC migraine, or angina pectoris. Specifically they are used to deliver
CC insulin or leuprolide, but many other suitable therapeutic agents are
CC disclosed, including genes or inhibitory nucleic acid, imaging agents and
CC antigens. (I) may also provide targeting to the GI tract. Other uses of
CC (I) are: (i) to determine the level of specified receptors in a sample
CC (in a binding assay); and (ii) to screen for molecules that bind (I).
CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
CC and in immunoassays. The antibodies are used to locate, detect and
CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
CC etc., also for peptide purification and immobilisation.
XX
SQ Sequence 708 AA;

Query Match 51.2%; Score 42; DB 20; Length 708;
Best Local Similarity 46.2%; Pred. No. 39;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GVKGTTVNANELP 15
|:|| |::: |:|
Db 524 gikgftissteip 536

RESULT 12
AAW74087
ID AAW74087 standard; peptide; 708 AA.
XX
AC AAW74087;
XX
DT 04-MAY-1999 (first entry)
XX
DE Gastro-intestinal transport receptor binding protein.
XX
KW Gastro-intestinal transport receptor; binding protein; hSI; HPT1;
KW D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
KW intestinal peptide-associated transporter; hypertension; diabetes;
KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KW therapeutic agent delivery; therapy.
XX
OS Homo sapiens.
XX
PN WO9851325-A2.
XX
PD 19-NOV-1998.
XX
PF 15-MAY-1998; 98WO-US10088.
XX
PR 15-MAY-1997; 97US-0046595.
XX
PA (CYTO-) CYTOGEN CORP.
PA (ELAN-) ELAN CORP PLC.
XX
PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
PI Omahony DJ, Patterson CA, Singleton J;
XX
DR WPI; 1999-009568/01.
XX
PT New proteins that bind specifically to receptors in the

PT gastro-intestinal tract and related nucleic acid - chimaeras and
PT antibodies, used to deliver therapeutic or diagnostic agents to, or
PT through, the gastrointestinal tract, e.g. insulin or leuprolide
XX
PS Claim 48; Page 163-164; 294pp; English.
XX
CC The invention relates to purified proteins (I) that bind
CC specifically to at least one of the gastro-intestinal (GI) tract
CC receptors human intestinal peptide-associated transporter (HPT1),
CC hPEPT1, D2H and human sucrose-isomaltase complex (hSI). (I) provide
CC active transport of therapeutic agents through human and animal GI tissue
CC (into the blood) for in vivo delivery, particularly for treatment or
CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
CC cancer, migraine, or angina pectoris. Specifically they are used to
CC deliver insulin or leuprolide, but many other suitable therapeutic agents
CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
CC and antigens. (I) may also provide targeting to the GI tract. Other uses
CC of (I) are: (i) to determine the level of specified receptors in a sample
CC (in a binding assay); and (ii) to screen for molecules that bind (I).
CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
CC and in immunoassays. The antibodies are used to locate, detect and
CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
CC etc., also for peptide purification and immobilisation.
CC Note: This sequence is claimed as being a nucleotide sequence.
XX
SQ Sequence 708 AA;

Query Match 51.2%; Score 42; DB 20; Length 708;
Best Local Similarity 46.2%; Pred. No. 39;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GVKGTTVNANELP 15
|:|| |::: |:|
Db 524 gikgftissteip 536

RESULT 13
AAW83394
ID AAW83394 standard; Protein; 708 AA.
XX
AC AAW83394;
XX
DT 23-FEB-1999 (first entry)
XX
DE Human protein-coupled peptide transporter.
XX
KW Human; protein-coupled peptide transporter; cell membrane;
KW blood-brain barrier; transportation.
XX
OS Homo sapiens.
XX
PN US5849525-A.
XX
PD 15-DEC-1998.
XX
PF 21-DEC-1995; 95US-0576165.
XX
PR 09-MAR-1994; 94US-0208645.
PR 21-DEC-1995; 95US-0576165.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
XX
PI Hediger M;
XX
DR WPI; 1999-069726/06.
DR N-PSDB; AAV72892.
XX
PT Nucleic acid encoding proton-coupled peptide transporter - providing
PT transport of peptides, their mimetics and drugs coupled to peptides
PT across cell membranes, including the blood-brain barrier
XX
PS Claim 2; Fig 2; 29pp; English.
XX
CC The present sequence represents human proton-coupled peptide transporter
CC (PCPT). Transforming cells with a nucleic acid molecule encoding PCPT
CC allows transport of peptides or their mimics across a cellular membrane.
CC Any di- or tri-peptide (and structurally similar compounds such as
CC beta-lactam antibiotics) can be transported into cells of the
CC gastrointestinal tract, brain, the blood-brain barrier, kidney and
CC liver. Chemicals (particularly therapeutic agents) can be coupled to
CC the peptides for delivery across membranes (with subsequent release of
CC active drug by enzymatic hydrolysis). Sequences antisense to the nucleic
CC acid molecule encoding PCPT may be used to inhibit PCPT expression.
XX
SQ Sequence 708 AA;

Query Match 51.2%; Score 42; DB 20; Length 708;
Best Local Similarity 46.2%; Pred. No. 39;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GVKGTTVNANELP 15
|:|| |::: |:|
Db 524 gikgftissteip 536

RESULT 14
AAE04789
ID AAE04789 standard; Protein; 605 AA.
XX
AC AAE04789;
XX
DT 10-SEP-2001 (first entry)
XX
DE Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCED1.
XX
KW Tomato; neoxanthin cleavage enzyme; LeNCED1; abscisic acid; ABA;
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;
KW plant growth protectant; herbicide.
XX
OS Lycopersicon esculentum.
XX
PN EP1116794-A2.
XX
PD 18-JUL-2001.
XX
PF 11-JAN-2001; 2001EP-0300218.
XX
PR 13-JAN-2000; 2000JP-0010056.
PR 11-JAN-2001; 2001JP-0003476.
XX
PA (RIKE) RIKEN KK.
XX
PI Iuchi S, Kobayashi M, Shinozaki K;
XX
DR WPI; 2001-400081/43.
DR N-PSDB; AAD09401.
XX
PT A DNA encoding a protein with a neoxanthin cleavage activity for
PT producing transgenic plants with improved or decreased stress tolerance
PT -
XX
PS Claim 3; Fig 2; 101pp; English.
XX
CC The invention relates to neoxanthin cleavage enzymes and their
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a
CC plant when expressed in a plant cell. The invention also relates to
CC methods for increasing or decreasing stress tolerance in a plant by
CC introducing the DNA into the plant, and a transgenic plant into which a
CC neoxanthin cleavage enzyme is introduced. The improvement of stress
CC tolerance in plants is useful, for example in plant breeding. Neoxanthin
CC cleavage enzyme genes are useful for producing transgenic plants. An arid
CC land can be improved by growing transformant weed for several years and

CC then removing the weed by specifically lowering stress tolerance in the
CC weed by inducing an inducible promoter. The present sequence is
CC Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCED1 protein
CC related to the invention.
XX
SQ Sequence 605 AA;

Query Match 50.0%; Score 41; DB 22; Length 605;
Best Local Similarity 53.8%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FGVKGTTVNANEL 14
:| || :|||:|
Db 589 ygfhgtfinandl 601

RESULT 15
AAB72308
ID AAB72308 standard; Protein; 605 AA.
XX
AC AAB72308;
XX
DT 16-MAY-2001 (first entry)
XX
DE Neoxanthin cleavage enzyme-like protein amino acid sequence.
XX
KW Defence-related signalling gene; sunflower; neoxanthin cleavage enzyme;
KW NCE; amino acid permease; AAP; glutamic acid rich protein; GRP;
KW pathogen resistance; abscisic acid metabolism.
XX
OS Lycopersicon esculentum.
XX
PN WO200112801-A2.
XX
PD 22-FEB-2001.
XX
PF 17-AUG-2000; 2000WO-US22961.
XX
PR 18-AUG-1999; 99US-0149656.
PR 23-MAY-2000; 2000US-0206405.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (CURA-) CURAGEN CORP.
XX
PI Bidney DL, Crasta OR, Hu X, Lu G;
XX
DR WPI; 2001-211215/21.
XX
PT Novel isolated defence-related signalling gene isolated from sunflower
PT encoding neoxanthin cleavage enzyme, amino acid permease or glutamic
PT acid-rich protein useful for increasing resistance of plant to a
PT pathogen -
XX
PS Disclosure; Fig 1; 135pp; English.
XX
CC This invention relates to defence-related signalling genes isolated from
CC the sunflower (Helianthus annuus). The genes encode a neoxanthin cleavage
CC enzyme (NCE), an amino acid permease (AAP) and a glutamic acid rich
CC protein (GRP). The signalling gene is useful for increasing the
CC resistance of a plant to a pathogen such as fungus, virus, bacterium,
CC nematode or insect (e.g. European corn borer), preferably
CC Sclerotinia spp., Phoma spp., or Phomopsis spp, by stably incorporating a
CC construct containing the gene into the genome of the plant. The gene is
CC useful for regulating gene expression in a plant, in response to a
CC stimulus such as infection with a pathogen, damage from a pathogen,
CC hydrogen peroxide, jasmonic acid, methyl jasmonate, salicylic acid,
CC oxalic acid or expression of a gene encoding oxalic acid oxidase. The
CC genes are also useful for stem-preferred regulation of gene expression in
CC a plant. The genes are useful in agriculture, particularly in the
CC breeding of crop plants with improved agronomic traits, for modifying
CC abscisic acid (ABA) metabolism and for modifying amino acid transport and
CC content in plants. The present sequence represents a neoxanthin cleavage
CC enzyme-like protein from Lycopersicon esculentum used in the
CC characterisation of sunflower NCE.
XX
SQ Sequence 605 AA;

Query Match 50.0%; Score 41; DB 22; Length 605;
Best Local Similarity 53.8%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FGVKGTTVNANEL 14
:| || :|||:|
Db 589 ygfhgtfinandl 601

Search completed: March 26, 2002, 13:38:47
Job time: 141 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:41:27 ; Search time 37.72 Seconds
(without alignments)
8.949 Million cell updates/sec

Title: US-09-709-201-97
Perfect score: 82
Sequence: 1 CFGVKGTTVNANELP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	51.2	708	2	US-08-576-165-2	Sequence 2, Appli
2	38.5	47.0	452	1	US-08-290-978A-5	Sequence 5, Appli
3	38.5	47.0	452	2	US-08-780-869-5	Sequence 5, Appli
4	38	46.3	322	2	US-08-286-819A-2	Sequence 2, Appli
5	38	46.3	322	3	US-08-980-357-2	Sequence 2, Appli
6	38	46.3	380	3	US-08-468-846-2	Sequence 2, Appli
7	38	46.3	380	4	US-08-915-096A-2	Sequence 2, Appli
8	38	46.3	2291	2	US-08-286-819A-29	Sequence 29, Appl
9	38	46.3	2291	3	US-08-980-357-29	Sequence 29, Appl
10	37	45.1	469	4	US-08-378-313-33	Sequence 33, Appl
11	36	43.9	916	1	US-08-188-228-48	Sequence 48, Appl
12	36	43.9	916	1	US-08-332-643-42	Sequence 42, Appl
13	36	43.9	916	1	US-08-332-638-48	Sequence 48, Appl
14	35.5	43.3	185	2	US-08-463-911-3	Sequence 3, Appli
15	35.5	43.3	236	4	US-09-140-804-6	Sequence 6, Appli
16	35	42.7	664	1	US-08-421-661-6	Sequence 6, Appli
17	35	42.7	992	4	US-08-890-865A-1	Sequence 1, Appli
18	34.5	42.1	200	4	US-09-282-146-2	Sequence 2, Appli
19	34	41.5	121	3	US-08-560-005-8	Sequence 8, Appli
20	34	41.5	121	4	US-09-418-540-8	Sequence 8, Appli
21	34	41.5	263	4	US-08-776-059-43	Sequence 43, Appl
22	34	41.5	264	4	US-08-776-059-33	Sequence 33, Appl
23	34	41.5	564	4	US-08-776-059-35	Sequence 35, Appl
24	34	41.5	953	2	US-08-500-857A-2	Sequence 2, Appli
25	34	41.5	1147	1	US-08-131-365B-38	Sequence 38, Appl
26	34	41.5	1147	2	US-08-668-123-38	Sequence 38, Appl
27	34	41.5	1297	4	US-09-540-245A-17	Sequence 17, Appl
28	34	41.5	2556	1	US-08-185-432-17	Sequence 17, Appl
29	34	41.5	2556	1	US-08-083-590A-20	Sequence 20, Appl
30	34	41.5	2556	3	US-08-532-384-20	Sequence 20, Appl
31	33.5	40.9	559	1	US-08-030-096-6	Sequence 6, Appli
32	33	40.2	31	2	US-08-023-980B-30	Sequence 30, Appl
33	33	40.2	31	2	US-08-486-953A-25	Sequence 25, Appl
34	33	40.2	79	1	US-08-154-916-12	Sequence 12, Appl
35	33	40.2	151	2	US-08-722-050-8	Sequence 8, Appli
36	33	40.2	237	4	US-08-861-774E-68	Sequence 68, Appl
37	33	40.2	317	6	5340934-11	Patent No. 5340934
38	33	40.2	369	1	US-07-854-596B-31	Sequence 31, Appl
39	33	40.2	385	2	US-08-892-715-2	Sequence 2, Appli
40	33	40.2	385	2	US-09-145-947-2	Sequence 2, Appli
41	33	40.2	403	2	US-08-533-669A-10	Sequence 10, Appl
42	33	40.2	403	2	US-08-607-509-2	Sequence 2, Appli
43	33	40.2	403	2	US-08-454-036-2	Sequence 2, Appli
44	33	40.2	403	2	US-08-634-642-2	Sequence 2, Appli
45	33	40.2	403	3	US-08-989-370-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-576-165-2
; Sequence 2, Application US/08576165
; Patent No. 5849525
; GENERAL INFORMATION:
; APPLICANT: HEDIGER, MATTHIAS
; TITLE OF INVENTION: COMPOSITIONS CORRESPONDING TO A
; TITLE OF INVENTION: PROTON-COUPLED PEPTIDE TRANSPORTER AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/576,165
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,645
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: JANIUK, ANTHONY J.
; REGISTRATION NUMBER: 29,809
; REFERENCE/DOCKET NUMBER: B0801/7022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 708 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-576-165-2

Query Match 51.2%; Score 42; DB 2; Length 708;
Best Local Similarity 46.2%; Pred. No. 11;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GVKGTTVNANELP 15

```
          |:|| |::: |:|
Db    524 GIKGFTISSTEIP 536


RESULT   2
US-08-290-978A-5
; Sequence 5, Application US/08290978A
; Patent No. 5624834
;  GENERAL INFORMATION:
;    APPLICANT:  KUSTERS-VAN SOMEREN, MARGO A.
;    APPLICANT:  MULLER, YVONNE
;    APPLICANT:  KESTER, HERMANUS C.M.
;    APPLICANT:  VISSER, JACOB
;    APPLICANT:  VAN OOYEN, ALBERT J.J.
;    APPLICANT:  ROLIN, CLAUS
;    TITLE OF INVENTION:  CLONING AND EXPRESSION OF THE
;    TITLE OF INVENTION:  EXO-POLYGALACTURONASE GENE FROM ASPERGILLUS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/290,978A
;      FILING DATE:  17-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  4615-0044.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  452 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-290-978A-5


  Query Match             47.0%;  Score 38.5;  DB 1;  Length 452;
  Best Local Similarity   53.3%;  Pred. No. 30;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       1 CFGVKGTTVNANELP 15
           |:| | ||:  || |
Db     365 CYGQKNTTL-CNEYP 378


RESULT   3
US-08-780-869-5
; Sequence 5, Application US/08780869
; Patent No. 5830737
;  GENERAL INFORMATION:
;    APPLICANT:  KUSTERS-VAN SOMEREN, MARGO A.
;    APPLICANT:  MULLER, YVONNE
;    APPLICANT:  KESTER, HERMANUS C.M.
;    APPLICANT:  VISSER, JACOB
;    APPLICANT:  VAN OOYEN, ALBERT J.J.
;    APPLICANT:  ROLIN, CLAUS
;    TITLE OF INVENTION:  CLONING AND EXPRESSION OF THE
;    TITLE OF INVENTION:  EXO-POLYGALACTURONASE GENE FROM ASPERGILLUS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/780,869
;      FILING DATE:  24-JAN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/290,978
;      FILING DATE:  17-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  4615-0044.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  452 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-780-869-5


  Query Match             47.0%;  Score 38.5;  DB 2;  Length 452;
  Best Local Similarity   53.3%;  Pred. No. 30;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       1 CFGVKGTTVNANELP 15
           |:| | ||:  || |
Db     365 CYGQKNTTL-CNEYP 378


RESULT   4
US-08-286-819A-2
; Sequence 2, Application US/08286819A
; Patent No. 5871910
;  GENERAL INFORMATION:
;    APPLICANT:  ARTHUR, MICHEL
;    APPLICANT:  DUKTA-MALEN, SYLVIE
;    APPLICANT:  MOLINAS, CATHERINE
;    APPLICANT:  COURVALIN, PATRICE
;    TITLE OF INVENTION:  POLYPEPTIDES IMPLICATED IN THE
;    TITLE OF INVENTION:  EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
;    TITLE OF INVENTION:  IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
;    TITLE OF INVENTION:  THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,819A
;      FILING DATE:  05-AUG-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/174,682
;      FILING DATE:  28-DEC-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/917,146
;      FILING DATE:  10-AUG-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR/91/00855
;      FILING DATE:  29-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9013579
;      FILING DATE:  31-OCT-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5871910man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  660-060-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  322 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-286-819A-2


  Query Match           46.3%;  Score 38;  DB 2;  Length 322;
  Best Local Similarity  63.6%;  Pred. No. 24;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNAN 12
           |||  | :|||
Db      26 FGVMATIINAN 36


RESULT   5
US-08-980-357-2
; Sequence 2, Application US/08980357
; Patent No. 6013508
;  GENERAL INFORMATION:
;    APPLICANT:  ARTHUR, MICHEL
;    APPLICANT:  DUKTA-MALEN, SYLVIE
;    APPLICANT:  MOLINAS, CATHERINE
;    APPLICANT:  COURVALIN, PATRICE
;    TITLE OF INVENTION:  POLYPEPTIDES IMPLICATED IN THE
;    TITLE OF INVENTION:  EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
;    TITLE OF INVENTION:  IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
;    TITLE OF INVENTION:  THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/980,357
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,819
;      FILING DATE:  05-AUG-1994
;      APPLICATION NUMBER:  US 08/174,682
;      FILING DATE:  28-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/917,146
;      FILING DATE:  10-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR/91/00855
;      FILING DATE:  29-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9013579
;      FILING DATE:  31-OCT-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 6013508man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  660-060-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  322 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-980-357-2


  Query Match           46.3%;  Score 38;  DB 3;  Length 322;
  Best Local Similarity  63.6%;  Pred. No. 24;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 FGVKGTTVNAN 12
           |||  | :|||
Db      26 FGVMATIINAN 36


RESULT   6
US-08-468-846-2
; Sequence 2, Application US/08468846
; Patent No. 6074839
;  GENERAL INFORMATION:
;    APPLICANT:  Meissner, Paul
;    APPLICANT:  Fuldner, Rebecca
;    APPLICANT:  Fei-Wei, Ying
;    APPLICANT:  Adams, Mark
;    TITLE OF INVENTION:  TRANSFORMING GROWTH FACTOR ALPHA HI
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;      ADDRESSEE:  STUART & OLSTEIN
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,846
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/208,008
```

```
;      FILING DATE:  08-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-465
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  380 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-468-846-2


  Query Match             46.3%;  Score 38;  DB 3;  Length 380;
  Best Local Similarity   50.0%;  Pred. No. 30;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CFGVKGTTVNANEL 14
           | | || ::| :||
Db      54 CPGGKGKSINCSEL 67


RESULT   7
US-08-915-096A-2
; Sequence 2, Application US/08915096A
; Patent No. 6265543
;  GENERAL INFORMATION:
;    APPLICANT:  Meissner, Paul S.
;    APPLICANT:  Fuldner, Rebecca A.
;    APPLICANT:  Adams, Mark D.
;    TITLE OF INVENTION:  Transforming Growth Factor Alpha HI
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,096A
;      FILING DATE:  20-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/468,846
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/208,008
;      FILING DATE:  08-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF110D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  380 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-915-096A-2


  Query Match             46.3%;  Score 38;  DB 4;  Length 380;
  Best Local Similarity   50.0%;  Pred. No. 30;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CFGVKGTTVNANEL 14
           | | || ::| :||
Db      54 CPGGKGKSINCSEL 67


RESULT   8
US-08-286-819A-29
; Sequence 29, Application US/08286819A
; Patent No. 5871910
;  GENERAL INFORMATION:
;    APPLICANT:  ARTHUR, MICHEL
;    APPLICANT:  DUKTA-MALEN, SYLVIE
;    APPLICANT:  MOLINAS, CATHERINE
;    APPLICANT:  COURVALIN, PATRICE
;    TITLE OF INVENTION:  POLYPEPTIDES IMPLICATED IN THE
;    TITLE OF INVENTION:  EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
;    TITLE OF INVENTION:  IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
;    TITLE OF INVENTION:  THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,819A
;      FILING DATE:  05-AUG-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/174,682
;      FILING DATE:  28-DEC-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/917,146
;      FILING DATE:  10-AUG-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR/91/00855
;      FILING DATE:  29-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9013579
;      FILING DATE:  31-OCT-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5871910man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  660-060-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2291 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-286-819A-29
```

```
  Query Match             46.3%;  Score 38;  DB 2;  Length 2291;
  Best Local Similarity   63.6%;  Pred. No. 2.6e+02;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 FGVKGTTVNAN 12
          |||  | :|||
Db   1138 FGVMATIINAN 1148


RESULT   9
US-08-980-357-29
; Sequence 29, Application US/08980357
; Patent No. 6013508
;  GENERAL INFORMATION:
;    APPLICANT:  ARTHUR, MICHEL
;    APPLICANT:  DUKTA-MALEN, SYLVIE
;    APPLICANT:  MOLINAS, CATHERINE
;    APPLICANT:  COURVALIN, PATRICE
;    TITLE OF INVENTION:  POLYPEPTIDES IMPLICATED IN THE
;    TITLE OF INVENTION:  EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
;    TITLE OF INVENTION:  IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
;    TITLE OF INVENTION:  THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/980,357
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,819
;      FILING DATE:  05-AUG-1994
;      APPLICATION NUMBER:  US 08/174,682
;      FILING DATE:  28-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/917,146
;      FILING DATE:  10-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR/91/00855
;      FILING DATE:  29-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9013579
;      FILING DATE:  31-OCT-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 6013508man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  660-060-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2291 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-980-357-29


  Query Match             46.3%;  Score 38;  DB 3;  Length 2291;
  Best Local Similarity   63.6%;  Pred. No. 2.6e+02;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 FGVKGTTVNAN 12
          |||  | :|||
Db   1138 FGVMATIINAN 1148


RESULT  10
US-08-378-313-33
; Sequence 33, Application US/08378313
; Patent No. 6207881
;  GENERAL INFORMATION:
;    APPLICANT:  THEOLOGIS, ATHANASIOS
;    APPLICANT:  SATO, TAKAHIDO
;    TITLE OF INVENTION:  CONTROL OF FRUIT RIPENING THROUGH
;    TITLE OF INVENTION:  GENETIC CONTROL OF ACC SYNTHASE SYNTHESIS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/378,313
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/862,493
;      FILING DATE:  02-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  29190-20002.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 856-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  469 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-378-313-33


  Query Match             45.1%;  Score 37;  DB 4;  Length 469;
  Best Local Similarity   77.8%;  Pred. No. 59;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GTTVNANEL 14
          |||:| |||
Db    205 GTTLNRNEL 213


RESULT  11
US-08-188-228-48
; Sequence 48, Application US/08188228
; Patent No. 5597725
;  GENERAL INFORMATION:
;    APPLICANT:  Suzuki, Shintaro
;    TITLE OF INVENTION:  CADHERIN MATERIALS AND METHODS
;    NUMBER OF SEQUENCES:  62
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Borun
```

```
;       STREET:  6300 Sears Tower, 233 S. Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60606
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/188,228
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/049,460
;       FILING DATE:  19 APR 1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/872,643
;       FILING DATE:  17 APR 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5597725and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  31340
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  48:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  916 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-188-228-48


  Query Match             43.9%;  Score 36;  DB 1;  Length 916;
  Best Local Similarity   42.9%;  Pred. No. 2e+02;
  Matches    6;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CFGVKGTTVNANEL 14
           | | |||    | :
Db      62 CVGTKGTQYETNSM 75


RESULT  12
US-08-332-643-42
; Sequence 42, Application US/08332643
; Patent No. 5639634
;  GENERAL INFORMATION:
;    APPLICANT:  Suzuki, Shintaro
;    TITLE OF INVENTION:  CADHERIN MATERIALS AND METHODS
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/332,643
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/872,643
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5639634and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30795
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  916 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-332-643-42


  Query Match             43.9%;  Score 36;  DB 1;  Length 916;
  Best Local Similarity   42.9%;  Pred. No. 2e+02;
  Matches    6;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CFGVKGTTVNANEL 14
           | | |||    | :
Db      62 CVGTKGTQYETNSM 75


RESULT  13
US-08-332-638-48
; Sequence 48, Application US/08332638
; Patent No. 5646250
;  GENERAL INFORMATION:
;    APPLICANT:  Suzuki, Shintaro
;    TITLE OF INVENTION:  CADHERIN MATERIALS AND METHODS
;    NUMBER OF SEQUENCES:  62
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Borun
;      STREET:  6300 Sears Tower, 233 S. Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/332,638
;      FILING DATE:  01-NOV-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/872,643
;      FILING DATE:  17 APR 1992
;      APPLICATION NUMBER:  US/08/049,460
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5646250and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  31340
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  916 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-332-638-48
```

```
  Query Match           43.9%;  Score 36;  DB 1;  Length 916;
  Best Local Similarity  42.9%;  Pred. No. 2e+02;
  Matches    6;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CFGVKGTTVNANEL 14
           | | |||    | :
Db      62 CVGTKGTQYETNSM 75


RESULT  14
US-08-463-911-3
; Sequence 3, Application US/08463911
; Patent No. 5869330
; GENERAL INFORMATION:
;    APPLICANT:  Scherer, Philipp E.
;    APPLICANT:  Lodish, Harvey F.
;    TITLE OF INVENTION:  A NOVEL SERUM PROTEIN PRODUCED
;    TITLE OF INVENTION:  EXCLUSIVELY IN ADIPOCYTES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,911
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI95-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  185 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-463-911-3


  Query Match           43.3%;  Score 35.5;  DB 2;  Length 185;
  Best Local Similarity  64.3%;  Pred. No. 37;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       5 KGTT---VNANELP 15
           |||:   | |||||
Db      57 KGTSAFAVKANELP 70


RESULT  15
US-09-140-804-6
; Sequence 6, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT: Humes, Jacqueline M.
;  TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
;  FILE REFERENCE: 97-49
;  CURRENT APPLICATION NUMBER: US/09/140,804
;  CURRENT FILING DATE: 1998-08-26
;  EARLIER APPLICATION NUMBER: 60/056,983
;  EARLIER FILING DATE: 1997-08-26
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 236
;   TYPE: PRT
;   ORGANISM: Tamias sibricus
US-09-140-804-6


  Query Match           43.3%;  Score 35.5;  DB 4;  Length 236;
  Best Local Similarity  64.3%;  Pred. No. 50;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       5 KGTT---VNANELP 15
           |||:   | |||||
Db     108 KGTSAFAVKANELP 121


Search completed: March 26, 2002, 13:41:28
Job time: 302 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:36:26 ; Search time 81.51 Seconds
(without alignments)
15.449 Million cell updates/sec

Title: US-09-709-201-93
Perfect score: 91
Sequence: 1 CTGSAAANYTTAVDRPN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	17	20	AAW95320	Costant and variab
2	82	90.1	29	20	AAW84462	Peptide Cp1A deriv
3	82	90.1	29	20	AAW84556	Peptide C.p.1A der
4	82	90.1	343	20	AAY56771	C. trachomatis ser
5	82	90.1	391	20	AAY35319	Chlamydia pneumoni
6	69	75.8	23	20	AAW84468	Peptide Cp1Aimp de
7	64	70.3	389	20	AAW98188	Chlamydia psittaci
8	55	60.4	10	20	AAW84466	Peptide CpVDI deri
9	55	60.4	10	20	AAW84553	Peptide C.pVDI der
10	49	53.8	100	20	AAW95295	Chlamydial major o
11	48	52.7	1241	22	AAM25606	Human protein sequ
12	48	52.7	1330	22	AAB65630	Novel protein kina
13	47	51.6	343	20	AAY56769	C. trachomatis ser
14	43	47.3	17	20	AAW95323	Costant and variab
15	43	47.3	250	21	AAG06126	Arabidopsis thalia
16	43	47.3	250	21	AAG46452	Arabidopsis thalia
17	43	47.3	263	21	AAG06125	Arabidopsis thalia
18	43	47.3	263	21	AAG46451	Arabidopsis thalia
19	41	45.1	345	22	AAB85128	Carica cysteine pr
20	41	45.1	477	22	AAB70765	Human betal-adreno
21	41	45.1	1736	22	AAB36932	Hepatitis C virus
22	41	45.1	2813	19	AAW54347	Canine von Willebr
23	41	45.1	2813	21	AAY70557	Canine von Willebr
24	41	45.1	3910	14	AAR38470	ALL-1 protein. Ho
25	41	45.1	3910	16	AAR66462	ALL-1 (acute lymph
26	41	45.1	3969	15	AAR52971	Product of the cDN
27	40	44.0	3011	16	AAR67588	Hepatitis C virus
28	39	42.9	205	22	AAG92148	C glutamicum prote
29	39	42.9	217	22	AAB79574	Corynebacterium gl
30	39	42.9	356	22	AAB99749	Oryza sativa perox
31	39	42.9	375	20	AAY30532	A G protein-couple
32	39	42.9	375	21	AAY71298	Human orphan G pro
33	39	42.9	375	21	AAB02832	Human G protein co
34	39	42.9	377	20	AAY30536	A G protein-couple
35	39	42.9	411	19	AAW80938	Human kidney lecit
36	39	42.9	414	19	AAW80941	Human kidney lecit
37	39	42.9	444	19	AAW80943	Human kidney lecit
38	38	41.8	158	21	AAY78913	Androgen independe
39	38	41.8	272	13	AAR29871	HCV NS4-NS5 peptid
40	38	41.8	374	22	AAG90207	C glutamicum prote
41	38	41.8	422	21	AAG28827	Arabidopsis thalia
42	38	41.8	422	21	AAG32115	Arabidopsis thalia
43	38	41.8	438	21	AAG51345	Arabidopsis thalia
44	38	41.8	459	19	AAW38456	Schizosaccharomyce
45	38	41.8	643	22	AAY97666	Zea mays ZmEIN3-1

ALIGNMENTS

RESULT 1
AAW95320
ID AAW95320 standard; Protein; 17 AA.
XX
AC AAW95320;
XX
DT 15-MAR-1999 (first entry)
XX
DE Costant and variable domain sequence of C. pneumoniae CPN90-105.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia pneumoniae.
XX
PN WO9850074-A2.
XX
PD 12-NOV-1998.
XX
PF 06-MAY-1998; 98WO-US09237.
XX
PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
PR 18-FEB-1998; 98US-0025176.
XX

PA (UYVA-) UNIV VANDERBILT.
XX
PI Mitchell WM, Stratton CW;
XX
DR WPI; 1999-059653/05.
XX
PT Composition with two agents effective against different stages of
PT chlamydial life cycle - comprises agent targetted against cryptic
PT phase, against elementary body phase, against replicating phase,
PT probenicid and antiporphyric
XX
PS Claim 4; Fig 3; 138pp; English.
XX
CC The invention relates to the diagnosis and management of infections by
CC Chlamydia species. The invention provides a composition that comprises
CC at least two agents, where each of the agents is effective against a
CC different phase of the chlamydial life cycle. The agents are selected
CC from: (a) agents targetted against cryptic phase of chlamydial life
CC cycle; (b) agents targetted against elementary body phase of chlamydial
CC life cycle; (c) agents targetted against replicating phase of chlamydial
CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC is used to elicit a protective immune response to Chlamydia infection in
CC an animal or human and is applied until the animal or human tests
CC negative for Chlamydia infection. It is also used to treat biological
CC material infected with Chlamydia. Diagnostic kits for antibody assays
CC against recombinant major outer membrane protein (MOMP), and for DNA
CC amplification assays for chlamydial genes, are used to diagnose disease,
CC e.g. autoimmune disease, an inflammatory disease or a disease that
CC occurs in an immuno-compromised individual, associated with Chlamydia
CC infection. The kits are used to detect chlamydial elementary bodies in a
CC sample. They are also used to monitor and/or modify the course of therapy
CC in a patient. The treatment reduces the acellular load of infectious
CC Ebstein Barr virus. The method is also used to treat porphyria, by
CC reducing the number of elementary bodies and applying a drug, e.g.
CC cimetidine, and antioxidants, to reduce the adverse effects associated
CC with porphyria. Sequences AAW95320 to AAW95323 represent constant and
CC variable domain sequences of various Chlamydia species.
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 91; DB 20; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGSAAANYTTAVDRPN 17
||||||||||||||||| |
Db 1 ctgsaaanyttavdrpn 17

RESULT 2
AAW84462
ID AAW84462 standard; peptide; 29 AA.
XX
AC AAW84462;
XX
DT 23-MAR-1999 (first entry)
XX
DE Peptide Cp1A derived from a major outer membrane protein.
XX
KW Variable domain; major outer membrane protein; MOMP;
KW Chlamydia; detection; infection; vaccine.
XX
OS Synthetic.
OS Chlamydia pneumoniae.
XX
PN WO9857981-A2.
XX
PD 23-DEC-1998.
XX
PF 15-JUN-1998; 98WO-IL00277.
XX
PR 19-JUN-1997; 97IL-0121114.
XX
PA (SAVY-) SAVYON DIAGNOSTICS LTD.
XX
PI Ohana B;
XX
DR WPI; 1999-080945/07.
XX
PT New peptides derived from Chlamydia pneumoniae MOMP protein - useful
PT to detect C. pneumoniae infection
XX
PS Claim 2; Page 53; 39pp; English.
XX
CC The present peptide is derived from the variable domain of the
CC major outer membrane protein (MOMP) of Chlamydia pneumoniae. The
CC peptide is able to react with antibodies formed during C. pneumoniae
CC infection and characterised by having essentially very low
CC cross-reactivity towards antibodies against other Chlamydia species.
CC A mixture of such peptides (see also AAW84462-68) is used to detect
CC C. pneumoniae infection, and in the preparation of vaccines.
XX
SQ Sequence 29 AA;

Query Match 90.1%; Score 82; DB 20; Length 29;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TGSAAANYTTAVDRPN 17
|||||||||||||||| |
Db 9 tgsaaanyttavdrpn 24

RESULT 3
AAW84556
ID AAW84556 standard; peptide; 29 AA.
XX
AC AAW84556;
XX
DT 26-MAR-1999 (first entry)
XX
DE Peptide C.p.1A derived from the major outer membrane protein.
XX
KW Variable domain; immunodominant; major outer membrane protein; MOMP;
KW anti-MOMP antibody; Chlamydia; vaccine; C. trachomatis.
XX
OS Chlamydia pneumoniae.
XX
PN WO9900414-A1.
XX
PD 07-JAN-1999.
XX
PF 15-JUN-1998; 98WO-IL00276.
XX
PR 19-JUN-1997; 97IL-0121115.
XX
PA (SAVY-) SAVYON DIAGNOSTICS LTD.
XX
PI Ohana B;
XX
DR WPI; 1999-095677/08.
XX
PT Chlamydia trachomatis specific peptides useful in diagnostic assays
PT - derived from major outer membrane protein variable domains and
PT useful in mixtures to detect infection with or immunise against all
PT serovars
XX
PS Example 1; Page 24; 78pp; English.
XX
CC The present sequence represents a peptide derived from variable
CC domain 1 (VDI) of the Chlamydia pneumoniae major outer membrane
CC protein (MOMP). The specification also describes C. trachomatis
CC MOMP derived peptides which have specificity only to C. trachomatis
CC anti-MOMP antibodies and are non-cross reactive with anti-MOMP

```
CC   antibodies of other Chlamydia species. Such peptides are useful to
CC   detect C. trachomatis infections in humans. Mixtures of MOMP peptide
CC   mixtures allow detection of and vaccination against all C. trachomatis
CC   serovars, which is not possible with existing MOMP-derived peptides
CC   for C. trachomatis-specific detection.
XX
SQ   Sequence   29 AA;


  Query Match            90.1%;  Score 82;  DB 20;  Length 29;
  Best Local Similarity  100.0%;  Pred. No. 8.2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||||||||||||||||
Db      9 tgsaaanyttavdrpn 24


RESULT   4
AAY56771
ID   AAY56771 standard; Protein; 343 AA.
XX
AC   AAY56771;
XX
DT   22-FEB-2000  (first entry)
XX
DE   C. trachomatis serovar HuPn MOMP sequence.
XX
KW   Major outer membrane protein; MOMP; Chlamydia; vaccine; immune response;
KW   cellular response; immunogen; Th1-like CD4 response; mucosal immunity.
XX
OS   Chlamydia trachomatis.
XX
PN   WO9951745-A2.
XX
PD   14-OCT-1999.
XX
PF   07-APR-1999;   99WO-CA00292.
XX
PR   07-APR-1998;   98US-0055765.
XX
PA   (UYMA-) UNIV MANITOBA.
XX
PI   Bruhnam RC;
XX
DR   WPI; 1999-620205/53.
XX
PT   Non-replicating vector encoding fragments of the outer membrane protein
PT   of Chlamydia, useful in vaccines and as immunogen -
XX
PS   Disclosure; Fig 10 A-F; 52pp; English.
XX
CC   The invention provides a non-replicating vector that comprises, linked
CC   to a promoter, a nucleotide sequence that encodes a region containing at
CC   least one of the conserved domains 2, 3 and 5 of a major outer membrane
CC   protein (MOMP) of a Chlamydia strain. The vector is used: (a) in
CC   vaccines to generate a protective immune response (mainly cellular)
CC   against MOMP, and (b) as immunogens to raise anti-MOMP antibodies, useful
CC   in standard immunoassays. Immunization with the vector induces a broad
CC   spectrum of immune responses, including Th1-like CD4 responses and
CC   mucosal immunity, providing significant protection against subsequent
CC   challenge. Sequences AAY56757-71 represent MOMP sequences from a variety
CC   of serovars of C. trachomatis.
XX
SQ   Sequence   343 AA;


  Query Match            90.1%;  Score 82;  DB 20;  Length 343;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||||||||||||||||
Db     90 tgsaaanyttavdrpn 105


RESULT   5
AAY35319
ID   AAY35319 standard; Protein; 391 AA.
XX
AC   AAY35319;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Chlamydia pneumoniae transmembrane protein sequence.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 1130-1131; Disclosure; 1912pp; English.
XX
CC   AAY34584-Y35879 represent the proteins encoded by all the open reading
CC   frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC   C. pneumoniae causes respiratory disease such as pneumonia and
CC   bronchitis and is thought to be a contributing factor in heart
CC   disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC   nodosum or pharyngitis. The polypeptides encoded by the open reading
CC   frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC   immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC   nucleotides sequences can also be used as immunogenic compositions,
CC   especially where the vector directs the expression of a neutralising
CC   epitope of C. pneumoniae.
XX
SQ   Sequence   391 AA;


  Query Match            90.1%;  Score 82;  DB 20;  Length 391;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||||||||||||||||
Db     92 tgsaaanyttavdrpn 107


RESULT   6
AAW84468
ID   AAW84468 standard; peptide; 23 AA.
XX
AC   AAW84468;
XX
DT   23-MAR-1999  (first entry)
XX
DE   Peptide Cp1Aimp derived from a major outer membrane protein.
XX
KW   Variable domain; major outer membrane protein; MOMP;
```

KW Chlamydia; detection; infection; vaccine.
XX
OS Synthetic.
OS Chlamydia pneumoniae.
XX
PN WO9857981-A2.
XX
PD 23-DEC-1998.
XX
PF 15-JUN-1998; 98WO-IL00277.
XX
PR 19-JUN-1997; 97IL-0121114.
XX
PA (SAVY-) SAVYON DIAGNOSTICS LTD.
XX
PI Ohana B;
XX
DR WPI; 1999-080945/07.
XX
PT New peptides derived from Chlamydia pneumoniae MOMP protein - useful
PT to detect C. pneumoniae infection
XX
PS Claim 2; Page 54; 39pp; English.
XX
CC The present peptide is derived from the variable domain of the
CC major outer membrane protein (MOMP) of Chlamydia pneumoniae. The
CC peptide is able to react with antibodies formed during C. pneumoniae
CC infection and characterised by having essentially very low
CC cross-reactivity towards antibodies against other Chlamydia species.
CC A mixture of such peptides (see also AAW84462-68) is used to detect
CC C. pneumoniae infection, and in the preparation of vaccines.
XX
SQ Sequence 23 AA;

Query Match 75.8%; Score 69; DB 20; Length 23;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TGSAAANYTTAVDR 15
||||||||||||||
Db 9 tgsaaanyttavdr 22

RESULT 7
AAW98188
ID AAW98188 standard; Protein; 389 AA.
XX
AC AAW98188;
XX
DT 05-JUL-1999 (first entry)
XX
DE Chlamydia psittaci major outer membrane protein.
XX
KW Major outer membrane protein; MOMP; psittacosis; infection;
KW vaccine; genetic immunisation.
XX
OS Chlamydia psittaci.
XX
PN WO9910005-A1.
XX
PD 04-MAR-1999.
XX
PF 28-AUG-1998; 98WO-US17943.
XX
PR 28-AUG-1997; 97US-0057147.
XX
PA (LOUU) UNIV LOUISIANA & AGRIC & MECH COLLEGE.
XX
PI Baghian A, Chouljenko VN, Kousoulas KG, Tully TN;
XX
DR WPI; 1999-254214/21.
DR N-PSDB; AAX25047.
XX
PT A new vaccine for Chlamydia psittaci infections
XX
PS Disclosure; Page 60-61; 72pp; English.
XX
CC The present sequence is the major outer membrane protein (MOMP)
CC of Chlamydia psittaci strain B577. A claimed MOMP polypeptide (see
CC AAW98184) comprises regions VD3 and VD4 of B577 MOMP, i.e. it lacks
CC regions VD1 and VD2. A claimed vaccine composition includes MOMP
CC polypeptide lacking VD1 and VD2, optionally fused to a maltose
CC binding protein. Also claimed are an isolated nucleic acid
CC encoding the polypeptide, a vector, and a method of preventing C.
CC psittaci infection by administering the vaccine containing the
CC MOMP polypeptide. Vectors encoding MOMP polypeptides lacking
CC regions VD1 and VD2 are useful for genetic vaccination. The
CC vaccines are used to prevent C. psittaci infection, especially in
CC birds.
XX
SQ Sequence 389 AA;

Query Match 70.3%; Score 64; DB 20; Length 389;
Best Local Similarity 75.0%; Pred. No. 0.012;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TGSAAANYTTAVDRPN 17
||:||||| | ||||
Db 90 tgtaaanyktptdrpn 105

RESULT 8
AAW84466
ID AAW84466 standard; peptide; 10 AA.
XX
AC AAW84466;
XX
DT 23-MAR-1999 (first entry)
XX
DE Peptide CpVDI derived from a major outer membrane protein.
XX
KW Variable domain; major outer membrane protein; MOMP;
KW Chlamydia; detection; infection; vaccine.
XX
OS Synthetic.
OS Chlamydia pneumoniae.
XX
PN WO9857981-A2.
XX
PD 23-DEC-1998
XX
PF 15-JUN-1998; 98WO-IL00277.
XX
PR 19-JUN-1997; 97IL-0121114.
XX
PA (SAVY-) SAVYON DIAGNOSTICS LTD.
XX
PI Ohana B;
XX
DR WPI; 1999-080945/07.
XX
PT New peptides derived from Chlamydia pneumoniae MOMP protein - useful
PT to detect C. pneumoniae infection
XX
PS Claim 2; Page 53; 39pp; English.
XX
CC The present peptide is derived from the variable domain of the
CC major outer membrane protein (MOMP) of Chlamydia pneumoniae. The
CC peptide is able to react with antibodies formed during C. pneumoniae
CC infection and characterised by having essentially very low
CC cross-reactivity towards antibodies against other Chlamydia species.
CC A mixture of such peptides (see also AAW84462-68) is used to detect
CC C. pneumoniae infection, and in the preparation of vaccines.
XX

```
SQ   Sequence   10 AA;

  Query Match            60.4%;  Score 55;  DB 20;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.0082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 NYTTAVDRPN 17
          ||||||||||
Db      1 nyttavdrpn 10


RESULT   9
AAW84553
ID   AAW84553 standard; peptide; 10 AA.
XX
AC   AAW84553;
XX
DT   26-MAR-1999  (first entry)
XX
DE   Peptide C.pVDI derived from the major outer membrane protein.
XX
KW   Variable domain; immunodominant; major outer membrane protein; MOMP;
KW   anti-MOMP antibody; Chlamydia; vaccine; C. trachomatis.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9900414-A1.
XX
PD   07-JAN-1999.
XX
PF   15-JUN-1998;   98WO-IL00276.
XX
PR   19-JUN-1997;   97IL-0121115.
XX
PA   (SAVY-) SAVYON DIAGNOSTICS LTD.
XX
PI   Ohana B;
XX
DR   WPI; 1999-095677/08.
XX
PT   Chlamydia trachomatis specific peptides useful in diagnostic assays
PT   - derived from major outer membrane protein variable domains and
PT   useful in mixtures to detect infection with or immunise against all
PT   serovars
XX
PS   Example 1; Page 21; 78pp; English.
XX
CC   The present sequence represents a peptide derived from variable
CC   domain 1 (VDI) of the Chlamydia pneumoniae major outer membrane
CC   protein (MOMP). The specification also describes C. trachomatis
CC   MOMP derived peptides which have specificity only to C. trachomatis
CC   anti-MOMP antibodies and are non-cross reactive with anti-MOMP
CC   antibodies of other Chlamydia species. Such peptides are useful to
CC   detect C. trachomatis infections in humans. Mixtures of MOMP peptide
CC   mixtures allow detection of and vaccination against all C. trachomatis
CC   serovars, which is not possible with existing MOMP-derived peptides
CC   for C. trachomatis-specific detection.
XX
SQ   Sequence   10 AA;

  Query Match            60.4%;  Score 55;  DB 20;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.0082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 NYTTAVDRPN 17
          ||||||||||
Db      1 nyttavdrpn 10


RESULT  10
AAW95295
ID   AAW95295 standard; Protein; 100 AA.
XX
AC   AAW95295;
XX
DT   15-MAR-1999  (first entry)
XX
DE   Chlamydial major outer membrane protein (MOMP) PN fragment.
XX
KW   Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW   antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW   major outer membrane protein; autoimmune; inflammatory; porphyria;
KW   Ebstein Barr virus; antioxidant.
XX
OS   Chlamydia sp.
XX
PN   WO9850074-A2.
XX
PD   12-NOV-1998.
XX
PF   06-MAY-1998;   98WO-US09237.
XX
PR   18-FEB-1998;   98US-0025521.
PR   06-MAY-1997;   97US-0045689.
PR   06-MAY-1997;   97US-0045739.
PR   06-MAY-1997;   97US-0045779.
PR   06-MAY-1997;   97US-0045780.
PR   06-MAY-1997;   97US-0045784.
PR   06-MAY-1997;   97US-0045787.
PR   14-AUG-1997;   97US-0911593.
PR   18-FEB-1998;   98US-0025174.
PR   18-FEB-1998;   98US-0025176.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Mitchell WM,  Stratton CW;
XX
DR   WPI; 1999-059653/05.
XX
PT   Composition with two agents effective against different stages of
PT   chlamydial life cycle - comprises agent targetted against cryptic
PT   phase, against elementary body phase, against replicating phase,
PT   probenicid and antiporphyric
XX
PS   Disclosure; Fig 1A; 138pp; English.
XX
CC   The invention relates to the diagnosis and management of infections by
CC   Chlamydia species. The invention provides a composition that comprises
CC   at least two agents, where each of the agents is effective against a
CC   different phase of the chlamydial life cycle. The agents are selected
CC   from: (a) agents targetted against cryptic phase of chlamydial life
CC   cycle; (b) agents targetted against elementary body phase of chlamydial
CC   life cycle; (c) agents targetted against replicating phase of chlamydial
CC   life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC   is used to elicit a protective immune response to Chlamydia infection in
CC   an animal or human and is applied until the animal or human tests
CC   negative for Chlamydia infection. It is also used to treat biological
CC   material infected with Chlamydia. Diagnostic kits for antibody assays
CC   against recombinant major outer membrane protein (MOMP), and for DNA
CC   amplification assays for chlamydial genes, are used to diagnose disease,
CC   e.g. autoimmune disease, an inflammatory disease or a disease that
CC   occurs in an immuno-compromised individual, associated with Chlamydia
CC   infection. The kits are used to detect chlamydial elementary bodies in a
CC   sample. They are also used to monitor and/or modify the course of therapy
CC   in a patient. The treatment reduces the acellular load of infectious
CC   Ebstein Barr virus. The method is also used to treat porphyria, by
CC   reducing the number of elementary bodies and applying a drug, e.g.
CC   cimetidine, and antioxidants, to reduce the adverse effects associated
CC   with porphyria. Sequences AAW95272 to AAW95319 represent peptide
CC   fragments of various Chlamydial MOMPs.
XX
SQ   Sequence   100 AA;
```

```
  Query Match           53.8%;  Score 49;  DB 20;  Length 100;
  Best Local Similarity  100.0%;  Pred. No. 0.9;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 YTTAVDRPN 17
           |||||||||
Db       1 yttavdrpn 9


RESULT  11
AAM25606
ID   AAM25606 standard; Protein; 1241 AA.
XX
AC   AAM25606;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:1121.
XX
KW   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW   antibacterial; endocrine; cardiant; central nervous system; virucide;
KW   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW   antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW   dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW   antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW   genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW   allergic rhinitis; diabetes; multiple sclerosis; depression;
KW   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW   neurological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200153455-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US35017.
XX
PR   23-DEC-1999;   99US-0471275.
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Drmanac RT;
XX
DR   WPI; 2001-457603/49.
DR   N-PSDB; AAH99547.
XX
PT   Isolated human polynucleotides encoding polypeptides, useful for the
PT   treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS   Claim 20; Page 232; 1217pp; English.
XX
CC   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC   AAM25963. The proteins can have activities based on the tissues and
CC   cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC   antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC   encoding them can be used in gene therapy, antisense therapy and vaccine
CC   production, The proteins and polynucleotides are useful for screening for
CC   agonists or antagonists of a protein and for the treatment and diagnosis
CC   of disorders associated with the activity of a protein e.g. inflammation,
CC   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC   infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC   osteoporosis, severe combined immunodeficiency, eczema, allergic
CC   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC   Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC   neurological disorders.
XX
SQ   Sequence   1241 AA;


  Query Match           52.7%;  Score 48;  DB 22;  Length 1241;
  Best Local Similarity  56.2%;  Pred. No. 18;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDRP 16
           ||||::| |  | | |
Db     619 ctgsssacyalatdlp 634


RESULT  12
AAB65630
ID   AAB65630 standard; Protein; 1330 AA.
XX
AC   AAB65630;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Novel protein kinase, SEQ ID NO: 156.
XX
KW   Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW   immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW   dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW   immune disorder; cardiovascular disease; neurodegenerative disease;
KW   cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW   inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
OS   Homo sapiens.
XX
PN   WO200073469-A2.
XX
PD   07-DEC-2000.
XX
PF   26-MAY-2000; 2000WO-US14842.
XX
PR   28-MAY-1999;   99US-0136503.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
DR   WPI; 2001-032161/04.
DR   N-PSDB; AAF44656.
XX
PT   Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT   treating immune-related diseases and disorders, cardiovascular disease,
PT   neurodegenerative diseases and/or cancers -
XX
PS   Claim 10; Fig 1; 310pp; English.
XX
CC   The present sequence is a novel protein kinase. The novel protein kinases
CC   and the nucleic acids that encode them may be used in the treatment and
CC   diagnosis of diseases associated with inappropriate kinase expression
CC   such as immune-related diseases and disorders, cardiovascular disease,
CC   neurodegenerative diseases and/or cancers. The nucleic acids and
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays. The kinase polypeptides may be used as antigens in the production
CC   of antibodies of kinase expression and activity. Anti-kinase antibodies
CC   and kinase antagonists may also be used to down regulate kinase
CC   expression and activity. Diseases related to kinase expression and
CC   activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC   disorders, complications of organ transplantation, myocardial infarction,
CC   immune disorders, cardiomyopathies, strokes, renal failure,
```

```
CC   oxidative-stress related disorders, chronic inflammatory bowel disease,
CC   chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC   osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC   reproductive disorders.
XX
SQ   Sequence   1330 AA;


  Query Match             52.7%;  Score 48;  DB 22;  Length 1330;
  Best Local Similarity   56.2%;  Pred. No. 20;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDRP 16
           ||||::| |  | | |
Db     708 ctgsssacyalatdlp 723


RESULT  13
AAY56769
ID   AAY56769 standard; Protein; 343 AA.
XX
AC   AAY56769;
XX
DT   22-FEB-2000  (first entry)
XX
DE   C. trachomatis serovar GPIC MOMP sequence.
XX
KW   Major outer membrane protein; MOMP; Chlamydia; vaccine; immune response;
KW   cellular response; immunogen; Th1-like CD4 response; mucosal immunity.
XX
OS   Chlamydia trachomatis.
XX
PN   WO9951745-A2.
XX
PD   14-OCT-1999.
XX
PF   07-APR-1999;   99WO-CA00292.
XX
PR   07-APR-1998;   98US-0055765.
XX
PA   (UYMA-) UNIV MANITOBA.
XX
PI   Bruhnam RC;
XX
DR   WPI; 1999-620205/53.
XX
PT   Non-replicating vector encoding fragments of the outer membrane protein
PT   of Chlamydia, useful in vaccines and as immunogen -
XX
PS   Disclosure; Fig 10 A-F; 52pp; English.
XX
CC   The invention provides a non-replicating vector that comprises, linked
CC   to a promoter, a nucleotide sequence that encodes a region containing at
CC   least one of the conserved domains 2, 3 and 5 of a major outer membrane
CC   protein (MOMP) of a Chlamydia strain. The vector is used: (a) in
CC   vaccines to generate a protective immune response (mainly cellular)
CC   against MOMP, and (b) as immunogens to raise anti-MOMP antibodies, useful
CC   in standard immunoassays. Immunization with the vector induces a broad
CC   spectrum of immune responses, including Th1-like CD4 responses and
CC   mucosal immunity, providing significant protection against subsequent
CC   challenge. Sequences AAY56757-71 represent MOMP sequences from a variety
CC   of serovars of C. trachomatis.
XX
SQ   Sequence   343 AA;


  Query Match             51.6%;  Score 47;  DB 20;  Length 343;
  Best Local Similarity   56.2%;  Pred. No. 7;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           ||:|||:: |  || |
Db      89 tgnaaadfktvadrnn 104


RESULT  14
AAW95323
ID   AAW95323 standard; Protein; 17 AA.
XX
AC   AAW95323;
XX
DT   15-MAR-1999  (first entry)
XX
DE   Costant and variable domain sequence of C. psitacci CPS92-106.
XX
KW   Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW   antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW   major outer membrane protein; autoimmune; inflammatory; porphyria;
KW   Ebstein Barr virus; antioxidant.
XX
OS   Chlamydia psitacci.
XX
PN   WO9850074-A2.
XX
PD   12-NOV-1998.
XX
PF   06-MAY-1998;   98WO-US09237.
XX
PR   18-FEB-1998;   98US-0025521.
PR   06-MAY-1997;   97US-0045689.
PR   06-MAY-1997;   97US-0045739.
PR   06-MAY-1997;   97US-0045779.
PR   06-MAY-1997;   97US-0045780.
PR   06-MAY-1997;   97US-0045784.
PR   06-MAY-1997;   97US-0045787.
PR   14-AUG-1997;   97US-0911593.
PR   18-FEB-1998;   98US-0025174.
PR   18-FEB-1998;   98US-0025176.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Mitchell WM;  Stratton CW;
XX
DR   WPI; 1999-059653/05.
XX
PT   Composition with two agents effective against different stages of
PT   chlamydial life cycle - comprises agent targetted against cryptic
PT   phase, against elementary body phase, against replicating phase,
PT   probenicid and antiporphyric
XX
PS   Claim 4; Fig 3; 138pp; English.
XX
CC   The invention relates to the diagnosis and management of infections by
CC   Chlamydia species. The invention provides a composition that comprises
CC   at least two agents, where each of the agents is effective against a
CC   different phase of the chlamydial life cycle. The agents are selected
CC   from: (a) agents targetted against cryptic phase of chlamydial life
CC   cycle; (b) agents targetted against elementary body phase of chlamydial
CC   life cycle; (c) agents targetted against replicating phase of chlamydial
CC   life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC   is used to elicit a protective immune response to Chlamydia infection in
CC   an animal or human and is applied until the animal or human tests
CC   negative for Chlamydia infection. It is also used to treat biological
CC   material infected with Chlamydia. Diagnostic kits for antibody assays
CC   against recombinant major outer membrane protein (MOMP), and for DNA
CC   amplification assays for chlamydial genes, are used to diagnose disease,
CC   e.g. autoimmune disease, an inflammatory disease or a disease that
CC   occurs in an immuno-compromised individual, associated with Chlamydia
CC   infection. The kits are used to detect chlamydial elementary bodies in a
CC   sample. They are also used to monitor and/or modify the course of therapy
CC   in a patient. The treatment reduces the acellular load of infectious
CC   Ebstein Barr virus. The method is also used to treat porphyria, by
CC   reducing the number of elementary bodies and applying a drug, e.g.
CC   cimetidine, and antioxidants, to reduce the adverse effects associated
CC   with porphyria. Sequences AAW95320 to AAW95323 represent constant and
CC   variable domain sequences of various Chlamydia species.
```

```
XX
SQ   Sequence   17 AA;


  Query Match             47.3%;  Score 43;  DB 20;  Length 17;
  Best Local Similarity   47.1%;  Pred. No. 1.4;
  Matches    8;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDRPN 17
           |    |:| | | || |
Db       1 casgtasnttvaadrsn 17


RESULT 15
AAG06126
ID   AAG06126 standard; Protein; 250 AA.
XX
AC   AAG06126;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 2786.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
```

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.3%; Score 43; DB 21; Length 250;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 SAAANYTTAVDR 15
           | | ||:||:||
Db      47 sralnystaidr 58
```

Search completed: March 26, 2002, 13:38:45
Job time: 139 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:36:26 ; Search time 37.72 Seconds
(without alignments)
10.142 Million cell updates/sec

Title: US-09-709-201-93
Perfect score: 91
Sequence: 1 CTGSAAANYTTAVDRPN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41	45.1	345	3	US-09-120-365-73	Sequence 73, Appl
2	41	45.1	345	4	US-09-515-039-73	Sequence 73, Appl
3	41	45.1	345	4	US-08-827-171B-7	Sequence 7, Appli
4	41	45.1	468	2	US-08-390-000A-7	Sequence 7, Appli
5	41	45.1	477	1	US-08-444-734A-4	Sequence 4, Appli
6	41	45.1	477	1	US-08-087-772A-16	Sequence 16, Appl
7	41	45.1	2813	3	US-08-896-449A-2	Sequence 2, Appli
8	41	45.1	2813	3	US-09-132-652-2	Sequence 2, Appli
9	41	45.1	3969	4	US-08-061-376-5	Sequence 5, Appli
10	40	44.0	472	1	US-08-194-338-6	Sequence 6, Appli
11	38	41.8	459	2	US-08-870-518-4	Sequence 4, Appli
12	38	41.8	918	4	US-09-041-886-11	Sequence 11, Appl
13	37	40.7	505	4	US-09-382-256-8	Sequence 8, Appli
14	37	40.7	505	4	US-09-395-115-8	Sequence 8, Appli
15	37	40.7	997	1	US-08-324-977-50	Sequence 50, Appl
16	37	40.7	997	2	US-08-384-616-50	Sequence 50, Appl
17	37	40.7	997	2	US-08-904-686A-50	Sequence 50, Appl
18	37	40.7	997	4	US-09-315-850-50	Sequence 50, Appl
19	37	40.7	2620	1	US-08-324-977-32	Sequence 32, Appl
20	37	40.7	2620	2	US-08-384-616-32	Sequence 32, Appl
21	37	40.7	2620	2	US-08-904-686A-32	Sequence 32, Appl
22	37	40.7	2620	4	US-09-315-850-32	Sequence 32, Appl
23	37	40.7	2621	1	US-08-324-977-36	Sequence 36, Appl
24	37	40.7	2621	2	US-08-384-616-36	Sequence 36, Appl
25	37	40.7	2621	2	US-08-904-686A-36	Sequence 36, Appl
26	37	40.7	2621	4	US-09-315-850-36	Sequence 36, Appl
27	37	40.7	3010	1	US-08-324-977-2	Sequence 2, Appli
28	37	40.7	3010	1	US-08-324-977-14	Sequence 14, Appl
29	37	40.7	3010	2	US-08-384-616-2	Sequence 2, Appli
30	37	40.7	3010	2	US-08-384-616-14	Sequence 14, Appl
31	37	40.7	3010	2	US-08-904-686A-2	Sequence 2, Appli
32	37	40.7	3010	2	US-08-904-686A-14	Sequence 14, Appl
33	37	40.7	3010	4	US-09-014-416-3	Sequence 3, Appli
34	37	40.7	3010	4	US-09-315-850-2	Sequence 2, Appli
35	37	40.7	3010	4	US-09-315-850-14	Sequence 14, Appl
36	36	39.6	120	1	US-08-539-304A-6	Sequence 6, Appli
37	36	39.6	304	4	US-09-088-651-2	Sequence 2, Appli
38	36	39.6	329	2	US-08-781-802-8	Sequence 8, Appli
39	36	39.6	329	4	US-08-694-078-8	Sequence 8, Appli
40	36	39.6	329	4	US-09-058-260-8	Sequence 8, Appli
41	36	39.6	366	3	US-08-945-056-6	Sequence 6, Appli
42	36	39.6	366	3	US-08-945-056-8	Sequence 8, Appli
43	36	39.6	374	2	US-08-915-107-2	Sequence 2, Appli
44	36	39.6	374	2	US-08-915-107-4	Sequence 4, Appli
45	36	39.6	374	4	US-09-273-613-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-120-365-73
; Sequence 73, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120,365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Papain
US-09-120-365-73

Query Match 45.1%; Score 41; DB 3; Length 345;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       2 TGSAAANYTT 11
           ||| | ||||
Db     109 TGSIAGNYTT 118
```

RESULT 2
US-09-515-039-73
; Sequence 73, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/515,039
; CURRENT FILING DATE: 2000-03-06
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Papain
US-09-515-039-73

```
  Query Match           45.1%;  Score 41;  DB 4;  Length 345;
  Best Local Similarity  80.0%;  Pred. No. 31;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTT 11
           ||| | ||||
Db     109 TGSIAGNYTT 118


RESULT   3
US-08-827-171B-7
; Sequence 7, Application US/08827171B
; Patent No. 6254869
;  GENERAL INFORMATION:
;    APPLICANT:  CAROLYN PETERSEN
;    APPLICANT:  JIN-XING HUANG
;    TITLE OF INVENTION:  CRYPTOPAIN VACCINES, ANTIBODIES, PROTEINS,
;    TITLE OF INVENTION:  PEPTIDES, DNA AND RNAS FOR PROPHYLAXIS,
;    TITLE OF INVENTION:  TREATMENT, DIAGNOSIS AND
;    TITLE OF INVENTION:  DETECTION OF
;    TITLE OF INVENTION:  CRYPTOSPORIDIUM PARVUM
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIK A
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Kb storage
;      COMPUTER:  PC
;      OPERATING SYSTEM:  WINDOWS
;      SOFTWARE:  Wordperfect 6.0a WINDOWS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,171B
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/014,233
;      FILING DATE:  March 27, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hana Verny
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  (HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-1677
;      TELEFAX:  (415) 324-1678
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  345 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Carica
US-08-827-171B-7


  Query Match           45.1%;  Score 41;  DB 4;  Length 345;
  Best Local Similarity  80.0%;  Pred. No. 31;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTT 11
           ||| | ||||
Db     109 TGSIAGNYTT 118


RESULT   4
US-08-390-000A-7
; Sequence 7, Application US/08390000A
; Patent No. 5985583
;  GENERAL INFORMATION:
;    APPLICANT:  Sealfon, Stuart C.
;    TITLE OF INVENTION:  Cloning and Expression of
;    TITLE OF INVENTION:  Gonadotropin-Releasing Hormone Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/390,000A
;      FILING DATE:  17-FEB-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  6923-052
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 790-9090
;      TELEFAX:  212 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-390-000A-7


  Query Match           45.1%;  Score 41;  DB 2;  Length 468;
  Best Local Similarity  43.8%;  Pred. No. 43;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDRP 16
           | | |||:  :::| |
Db     442 CNGGAAADSDSSLDEP 457


RESULT   5
US-08-444-734A-4
; Sequence 4, Application US/08444734A
; Patent No. 5610282
;  GENERAL INFORMATION:
;    APPLICANT:  Sibley, David R.
;    APPLICANT:  Monsma, Frederick J.
;    APPLICANT:  Mahan, Lawrence C.
;    APPLICANT:  McVittie, Loris D.
;    TITLE OF INVENTION:  cDNA encoding the rat D1 dopamine
;    TITLE OF INVENTION:  receptor linked to adnylyl cyclase activation and
;    TITLE OF INVENTION:  expression of the receptor protein in plasmid-transfected
;    TITLE OF INVENTION:  cell lines
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson and Bear
;      STREET:  620 Newport Center Drive, Sixteenth Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/444,734A
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/029,917
;       FILING DATE:  03-MAR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/548,714
;       FILING DATE:  06-JUL-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Altman, Daniel E.
;       REGISTRATION NUMBER:  34,115
;       REFERENCE/DOCKET NUMBER:  NIH065.001FW1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (714) 760-0404
;       TELEFAX:  (714) 760-9502
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  477 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FRAGMENT TYPE:  internal
US-08-444-734A-4


  Query Match             45.1%;  Score 41;  DB 1;  Length 477;
  Best Local Similarity   43.8%;  Pred. No. 44;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CTGSAAANYTTAVDRP 16
          | | |||:  :::| |
Db    451 CNGGAAADSDSSLDEP 466


RESULT   6
US-08-087-772A-16
; Sequence 16, Application US/08087772A
; Patent No. 5691155
;  GENERAL INFORMATION:
;    APPLICANT:  Nahmias, Clara
;    APPLICANT:  Emorine, Jean L.
;    APPLICANT:  Strosberg, Donny A.
;    TITLE OF INVENTION:  Nucleotide Sequences Encoding the Murine
;    TITLE OF INVENTION:  Beta3-Adrenergic Receptor and Their Applications
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bell, Seltzer, Park & Gibson
;      STREET:  Post Office Drawer 34009
;      CITY:  Charlotte
;      STATE:  No. 5691155th Carolina
;      COUNTRY:  USA
;      ZIP:  28234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/087,772A
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Linker, Raymond O.
;      REGISTRATION NUMBER:  26,419
;      REFERENCE/DOCKET NUMBER:  3339-195
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-881-3140
;      TELEFAX:  919-881-3175
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  477 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-087-772A-16


  Query Match             45.1%;  Score 41;  DB 1;  Length 477;
  Best Local Similarity   43.8%;  Pred. No. 44;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CTGSAAANYTTAVDRP 16
          | | |||:  :::| |
Db    451 CNGGAAADSDSSLDEP 466


RESULT   7
US-08-896-449A-2
; Sequence 2, Application US/08896449A
; Patent No. 6040143
;  GENERAL INFORMATION:
;    APPLICANT:  Venta, Patrick J
;    APPLICANT:  Yuzbasiyan-Gurkan, Vilma
;    APPLICANT:  Schall, William D
;    APPLICANT:  Brewer, George J
;    TITLE OF INVENTION:  DNA ENCODING CANINE VON WILLEBRAND
;    TITLE OF INVENTION:  FACTOR AND METHODS OF USE
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Harness, Dickey & Pierce, P.L.C.
;      STREET:  5445 Corporate Drive
;      CITY:  Troy
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  48098
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/896,449A
;      FILING DATE:  18-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, DeAnn F.
;      REFERENCE/DOCKET NUMBER:  2115-001226
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  248-641-1600
;      TELEFAX:  248-641-0270
;      TELEX:  287637
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2813 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-896-449A-2


  Query Match             45.1%;  Score 41;  DB 3;  Length 2813;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+02;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 CTGSAAANYTTAVDR 15
          |  || |||  || |
Db    621 CLCSAVANYAAAVAR 635
```

```
RESULT   8
US-09-132-652-2
; Sequence 2, Application US/09132652
; Patent No. 6074832
; GENERAL INFORMATION:
;  APPLICANT: Venta, Patrick J
;  APPLICANT: Yuzbasiyan-Gurkan, Vilma
;  APPLICANT: Schall, William D
;  APPLICANT: Brewer, George J
;  APPLICANT: Duffendeck, John
;  TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS
;  TITLE OF INVENTION: OF USE
;  FILE REFERENCE: 2115S-001226CPB
;  CURRENT APPLICATION NUMBER: US/09/132,652
;  CURRENT FILING DATE: 1998-08-11
;  EARLIER APPLICATION NUMBER: 08/896,449
;  EARLIER FILING DATE: 1997-07-18
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;    LENGTH: 2813
;    TYPE: PRT
;    ORGANISM: Canis familiaris
US-09-132-652-2


  Query Match             45.1%;  Score 41;  DB 3;  Length 2813;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+02;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDR 15
           |  || |||  || |
Db     621 CLCSAVANYAAAVAR 635


RESULT   9
US-08-061-376-5
; Sequence 5, Application US/08061376
; Patent No. 6175000
;  GENERAL INFORMATION:
;    APPLICANT:  Evans, Glen A.
;    APPLICANT:  Djabali, Malek
;    APPLICANT:  Selleri, Licia
;    APPLICANT:  Parry, Pauline
;    TITLE OF INVENTION:  CHARACTERIZATION OF A CHROMOSOME 11Q23
;    TITLE OF INVENTION:  TRANSLOCATION BREAKPOINT ASSOCIATED WITH ACUTE LEUKEMIAS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/061,376
;      FILING DATE:  13-MAY-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P41 9387
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)546-4737
;      TELEFAX:  (619)546-9392
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3969 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-061-376-5


  Query Match             45.1%;  Score 41;  DB 4;  Length 3969;
  Best Local Similarity   43.8%;  Pred. No. 4.6e+02;
  Matches    7;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           :|   :|:|  || ||
Db    3475 SGPQVSNFTQTVDAPN 3490


RESULT  10
US-08-194-338-6
; Sequence 6, Application US/08194338
; Patent No. 5474898
;  GENERAL INFORMATION:
;    APPLICANT:  Venter, John C.
;    APPLICANT:  Fraser, Claire M.
;    APPLICANT:  McCombie, William R.
;    TITLE OF INVENTION:  OCTOPAMINE RECEPTOR
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, olson and Bear
;      STREET:  620 Newport Center Drive, Sixteenth Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/194,338
;      FILING DATE:  08-FEB-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/676,174
;      FILING DATE:  28-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A.
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH101.001DV1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  472 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  internal
US-08-194-338-6


  Query Match             44.0%;  Score 40;  DB 1;  Length 472;
  Best Local Similarity   43.8%;  Pred. No. 64;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDRP 16
           | | |||:  :::| |
Db     446 CNGCAAADSDSSLDEP 461
```

```
RESULT  11
US-08-870-518-4
; Sequence 4, Application US/08870518
; Patent No. 5925566
;  GENERAL INFORMATION:
;    APPLICANT:  Davis, Roger J.
;    APPLICANT:  Galcheva-Gargova, Zoya
;    TITLE OF INVENTION:  NON-ACTIVATED RECEPTOR COMPLEX
;    TITLE OF INVENTION:  PROTEINS AND USES THEREOF
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/870,518
;      FILING DATE:  06-JUN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/019,219
;      FILING DATE:  06-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, Peter J.
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  04020/102001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  459 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-870-518-4


  Query Match             41.8%;  Score 38;  DB 2;  Length 459;
  Best Local Similarity   61.5%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GSAAANYTTAVDR 15
           |:|| | :|| ||
Db      14 GNAAQNVSTAEDR 26


RESULT  12
US-09-041-886-11
; Sequence 11, Application US/09041886
; Patent No. 6235872
;  GENERAL INFORMATION:
;    APPLICANT:  Bredesen, Dale E.
;    APPLICANT:  Rabizadeh, Sharroz
;    TITLE OF INVENTION:  Proapoptotic Peptides, Dependence
;    TITLE OF INVENTION:  Polypeptides and Methods of Use
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/041,886
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LJ 2626
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  918 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-041-886-11


  Query Match             41.8%;  Score 38;  DB 4;  Length 918;
  Best Local Similarity   41.2%;  Pred. No. 2.8e+02;
  Matches    7;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDRPN 17
             |:|||||  :  :: |:
Db     321 CSGSAAAGSSGTLELPS 337


RESULT  13
US-09-382-256-8
; Sequence 8, Application US/09382256A
; Patent No. 6207814
;  GENERAL INFORMATION:
;       APPLICANT: MIYAZONO, Kohei
;                  TEN DIJKE, Peter
;                  FRANZEN, Petra
;                  YAMASHITA, Hidetoshi
;                  HELDIN, Carl-Henrik
;       TITLE OF INVENTION: ACTIVIN RECEPTOR LIKE KINASES, PROTEINS
;                           HAVING SERINE THREONINE KINASE DOMAINS,
;                           AND THEIR USE
;       NUMBER OF SEQUENCES: 29
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fulbright & Jaworski L.L.P.
;            STREET: 666 Fifth Avenue
;            CITY: New York City
;            STATE: New York
;            COUNTRY: USA
;            ZIP: 10103
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb
;            COMPUTER: IBM PS/2
;            OPERATING SYSTEM: PC-DOS
;            SOFTWARE: Wordperfect
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/382.256A
;            FILING DATE: 24-Aug-1999
;            CLASSIFICATION: 514
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/GB93/02367
;            FILING DATE: No. 6207814ember 17, 1993
;            APPLICATION NUMBER: GB 9224057.1
;            FILING DATE: No. 6207814ember 17, 1992
;            APPLICATION NUMBER: GB 9304677.9
;            FILING DATE: March 8, 1993
;            APPLICATION NUMBER: GB 9304680.3
;            FILING DATE: March 8, 1993
;            APPLICATION NUMBER: 9311047.6
```

```
;             FILING DATE: May 28, 1993
;             APPLICATION NUMBER: 9313763.6
;             FILING DATE: July 2, 1993
;             APPLICATION NUMBER: 9316099.2
;             FILING DATE: August 3, 1993
;             APPLICATION NUMBER: 321344.5
;             FILING DATE: October 15, 1993
;       ATTORNEY/AGENT INFORMATION:
;             NAME: No. 6207814man D. Hanson
;             REGISTRATION NUMBER: 30,946
;             REFERENCE/DOCKET NUMBER: LUD 5298.1
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 318-3000
;             TELEFAX: (212) 752-5958
;   INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 505 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-382-256-8


  Query Match             40.7%;  Score 37;  DB 4;  Length 505;
  Best Local Similarity   50.0%;  Pred. No. 2.1e+02;
  Matches    7;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVD 14
           ||    ||||   |
Db      36 CTSCLQANYTCETD 49


RESULT  14
US-09-395-115-8
; Sequence 8, Application US/09395115
; Patent No. 6271365
;  GENERAL INFORMATION:
;    APPLICANT:  Miyazono, Kohei; Dijke, Peter Ten;
;    APPLICANT:  Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik
;    TITLE OF INVENTION:  Activin Receptor-Like Kinase, Proteins
;    TITLE OF INVENTION:  Having Serine Threonine Kinase Domains And Their Use
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/395,115
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/436,265
;      FILING DATE:  30-October-1995
;      APPLICATION NUMBER:  PCT/GB93/02367
;      FILING DATE:  17-No. 6271365ember-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  9224057.1
;      FILING DATE:  17-No. 6271365ember-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  9304677.9
;      FILING DATE:  8-March-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  9304680.3
;      FILING DATE:  8-March-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  9311047.6
;      FILING DATE:  28-May-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  9313763.6
;      FILING DATE:  2-July-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  9136099.2
;      FILING DATE:  3-August-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  9321344.5
;      FILING DATE:  15-October-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kohlei, Vineet
;      REGISTRATION NUMBER:  37,003
;      REFERENCE/DOCKET NUMBER:  LUD 5298
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  505 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-395-115-8


  Query Match             40.7%;  Score 37;  DB 4;  Length 505;
  Best Local Similarity   50.0%;  Pred. No. 2.1e+02;
  Matches    7;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVD 14
           ||    ||||   |
Db      36 CTSCLQANYTCETD 49


RESULT  15
US-08-324-977-50
; Sequence 50, Application US/08324977
; Patent No. 5747339
;  GENERAL INFORMATION:
;    APPLICANT:  OKAYAMA, Hiroto
;    APPLICANT:  FUKE, Isao
;    APPLICANT:  MORI, Chisato
;    APPLICANT:  TAKAMIZAWA, Akahisa
;    APPLICANT:  YOSHIDA, Iwao
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOMIC
;    TITLE OF INVENTION:  cDNA AND ANTIGEN POLYPEPTIDE
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Armstrong, Westerman, Hattori, McLeland &
;      ADDRESSEE:  Naughton
;      STREET:  1725 K St. N.W. Suite 1000
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in, 1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS, Version 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/324,977
;      FILING DATE:  18-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-167466
;      FILING DATE:  25-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-230921
;      FILING DATE:  31-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-305605
```

```
;      FILING DATE:  09-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/099,706
;      FILING DATE:  30-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/769,996
;      FILING DATE:  02-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/635,451
;      FILING DATE:  28-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens-Smith, Theresa M.
;      REGISTRATION NUMBER:  36,281
;      REFERENCE/DOCKET NUMBER:  900703D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 659-2930
;      TELEFAX:  (202) 887-0357
;      TELEX:  440142
;  INFORMATION FOR SEQ ID NO:  50:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  997 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-324-977-50


  Query Match             40.7%;  Score 37;  DB 1;  Length 997;
  Best Local Similarity   53.3%;  Pred. No. 4.5e+02;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDR 15
           || | | ||: |: |
Db      57 CTPSPAPNYSRALWR 71


Search completed: March 26, 2002, 13:41:27
Job time: 301 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:36:26 ; Search time 42.75 Seconds
(without alignments)
30.292 Million cell updates/sec

Title: US-09-709-201-93
Perfect score: 91
Sequence: 1 CTGSAAANYTTAVDRPN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	82	90.1	389	2	A43587	major outer membra
2	82	90.1	389	2	D86577	major outer membra
3	78	85.7	389	2	I40864	major outer membra
4	78	85.7	389	2	I40739	major outer membra
5	64	70.3	389	1	MMCWP3	major outer membra
6	48	52.7	662	2	T17211	hypothetical prote
7	47	51.6	56	2	D71349	probable ribosomal
8	47	51.6	389	2	A60109	major outer membra
9	46	50.5	168	2	A84156	single-strand DNA-
10	46	50.5	1325	2	T25753	hypothetical prote
11	45	49.5	3300	2	D70575	probable PPE prote
12	44	48.4	584	1	I39710	cellulose biosynth
13	43	47.3	429	2	T01009	hypothetical prote
14	42	46.2	306	2	S59540	heat shock transcr
15	42	46.2	422	2	D84403	dihydroorotase [im
16	42	46.2	1005	2	C71513	hypothetical prote
17	41	45.1	229	2	E70978	hypothetical prote
18	41	45.1	273	2	G81952	HemK protein NMA03
19	41	45.1	310	2	JC7275	acid nuclease Le1
20	41	45.1	345	1	PPPA	papain (EC 3.4.22.
21	41	45.1	423	2	E81010	hemK protein NMB20
22	41	45.1	431	2	JW0098	carbazole dioxygen
23	41	45.1	477	1	QRHUB1	beta-1-adrenergic
24	41	45.1	480	2	I53053	beta 1 adrenergic
25	41	45.1	693	2	JN0673	ubiquitin-like fus
26	41	45.1	3968	2	A44265	trithorax homolog
27	40	44.0	87	1	BXSMAC	antibacterial subs
28	40	44.0	280	2	G36808	hypothetical prote
29	40	44.0	337	2	B84335	hypothetical prote
30	40	44.0	375	2	E83163	hypothetical prote
31	40	44.0	395	2	T44512	hypothetical prote
32	40	44.0	422	2	G84302	hypothetical prote
33	39.5	43.4	268	2	H83051	dihydrodipicolinat
34	39	42.9	375	2	JC7287	G-protein coupled
35	39	42.9	414	2	F64409	hypothetical prote
36	39	42.9	473	2	C70915	hypothetical prote
37	39	42.9	2037	2	S37178	fatty-acyl-CoA syn
38	38.5	42.3	227	2	T45209	hypothetical prote
39	38.5	42.3	286	2	C70953	probable lpqZ prot
40	38.5	42.3	430	2	T45151	dihydroorotase (EC
41	38.5	42.3	454	2	A69017	dihydroorotase - M
42	38	41.8	116	2	H64226	hypothetical prote
43	38	41.8	121	2	S71007	hypothetical prote
44	38	41.8	125	2	T39097	very hypothetical
45	38	41.8	259	2	C81081	ferredoxin--NADP r

ALIGNMENTS

RESULT 1
A43587
major outer membrane protein, porin CP0051 precursor [imported] - Chlamydophila pneum
N;Alternate names: MOMP
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 11-May-2000
C;Accession: A43587; A49751; A49216; G72044; F81619
R;Perez Melgosa, M.; Kuo, C.C.; Campbell, L.A.
Infect. Immun. 59, 2195-2199, 1991
A;Title: Sequence analysis of the major outer membrane protein gene of Chlamydia pneu
A;Reference number: A43587; MUID:91244474
A;Accession: A43587
A;Molecule type: DNA
A;Residues: 1-389 <PER>
A;Cross-references: GB:M69230; NID:g144540; PIDN:AAA73071.1; PID:g144541
R;Carter, M.W.; Al-Mahdawi, S.A.H.; Giles, I.G.; Treharne, J.D.; Ward, M.E.; Clarke,
J. Gen. Microbiol. 137, 465-475, 1991
A;Title: Nucleotide sequence and taxonomic value of the major outer membrane protein
A;Reference number: A49751; MUID:91237311
A;Accession: A49751
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <CAR>
A;Cross-references: GB:M64064; GB:M34942; NID:g144534; PIDN:AAA23143.1; PID:g144535
A;Note: isolate IOL-207
R;Gaydos, C.A.; Quinn, T.C.; Bobo, L.D.; Eiden, J.J.
Infect. Immun. 60, 5319-5323, 1992
A;Title: Similarity of Chlamydia pneumoniae strains in the variable domain IV region
A;Reference number: A49216; MUID:93084388
A;Accession: A49216
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 297-352 <GAY>
A;Cross-references: GB:S50607; NID:g260972; PIDN:AAB24363.1; PID:g260973
A;Note: sequence extracted from NCBI backbone (NCBIN:120604, NCBIP:120605)
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: G72044
A;Molecule type: DNA
A;Residues: 1-389 <ARN>
A;Cross-references: GB:AE001652; GB:AE001363; NID:g4376997; PIDN:AAD18834.1; PID:g437
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255
A;Accession: F81619
A;Status: preliminary
A;Molecule type: DNA

```
A;Residues: 1-389 <REA>
A;Cross-references: GB:AE002168; GB:AE002161; NID:g7188982; PIDN:AAF37944.1; PID:g718899
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: ompA; CP0051
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match           90.1%;  Score 82;  DB 2;  Length 389;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           ||||||||||||||||
Db      90 TGSAAANYTTAVDRPN 105


RESULT   2
D86577
major outer membrane protein [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: D86577
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: D86577
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <STO>
A;Cross-references: GB:BA000008; NID:g8979067; PIDN:BAA98902.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: ompA
C;Superfamily: Chlamydia major outer membrane protein


  Query Match           90.1%;  Score 82;  DB 2;  Length 389;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           ||||||||||||||||
Db      90 TGSAAANYTTAVDRPN 105


RESULT   3
I40864
major outer membrane protein - Chlamydophila psittaci
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
C;Accession: I40864; S33465
R;Girjes, A.A.; Carrick, F.N.; Lavin, M.F.
Gene 138, 139-142, 1994
A;Title: Remarkable sequence relatedness in the DNA encoding the major outer membrane pr
A;Reference number: I40864; MUID:94171025
A;Accession: I40864
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-389 <RES>
A;Cross-references: EMBL:X72023; NID:g313844; PIDN:CAA50906.1; PID:g313845
C;Superfamily: Chlamydia major outer membrane protein


  Query Match           85.7%;  Score 78;  DB 2;  Length 389;
  Best Local Similarity  93.8%;  Pred. No. 1.9e-05;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           |||| ||||||||||
Db      90 TGSATANYTTAVDRPN 105


RESULT   4
I40739
major outer membrane protein precursor - Chlamydophila pneumoniae (strain equine/N16)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
A;Variety: strain equine/N16
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Apr-2000
C;Accession: I40739
R;Storey, C.; Lusher, M.; Yates, P.; Richmond, S.
J. Gen. Microbiol. 139, 2621-2626, 1993
A;Title: Evidence for Chlamydia pneumoniae of non-human origin.
A;Reference number: I40739; MUID:94103736
A;Accession: I40739
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-389 <STO>
A;Cross-references: GB:L04982; NID:g289840; PIDN:AAA17397.1; PID:g289841
C;Comment: On the basis of the major outer membrane protein the authors classified th
 the sequence of the genome strain CWL029 and strain strain IOL-207. See PIR:A43587.
C;Genetics:
A;Gene: momp
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match           85.7%;  Score 78;  DB 2;  Length 389;
  Best Local Similarity  93.8%;  Pred. No. 1.9e-05;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           |||| ||||||||||
Db      90 TGSATANYTTAVDRPN 105


RESULT   5
MMCWP3
major outer membrane protein precursor - Chlamydophila psittaci (strain S26/3)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 31-Mar-2000
C;Accession: S08770
R;Herring, A.J.; Tan, T.W.; Baxter, S.; Inglis, N.F.; Dunbar, S.
FEMS Microbiol. Lett. 65, 153-158, 1989
A;Title: Sequence analysis of the major outer membrane protein gene of an ovine abort
A;Reference number: S08770
A;Accession: S08770
A;Molecule type: DNA
A;Residues: 1-389 <HER>
A;Cross-references: EMBL:X51859; NID:g40600; PIDN:CAA36152.1; PID:g40601
C;Superfamily: Chlamydia major outer membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match           70.3%;  Score 64;  DB 1;  Length 389;
  Best Local Similarity  75.0%;  Pred. No. 0.0044;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           ||:|||| |  ||||
Db      90 TGTAAANYKTPTDRPN 105


RESULT   6
T17211
hypothetical protein DKFZp434O051.1 - human
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17211
```

```
R;Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18723
A;Accession: T17211
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-662 <POU>
A;Cross-references: EMBL:AL117400
A;Experimental source: adult testis; clone DKFZp434O051
C;Genetics:
A;Note: DKFZp434O051.1


  Query Match             52.7%;  Score 48;  DB 2;  Length 662;
  Best Local Similarity   56.2%;  Pred. No. 3.8;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CTGSAAANYTTAVDRP 16
          ||||::| |  | | |
Db    251 CTGSSSACYALATDLP 266


RESULT   7
D71349
probable ribosomal protein L33 (rpmG) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 13-Aug-1999
C;Accession: D71349
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770
A;Accession: D71349
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-56 <COL>
A;Cross-references: GB:AE001205; GB:AE000520; NID:g3322501; PIDN:AAC65222.1; PID:g33225
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0234
C;Superfamily: Escherichia coli ribosomal protein L33


  Query Match             51.6%;  Score 47;  DB 2;  Length 56;
  Best Local Similarity   52.9%;  Pred. No. 0.45;
  Matches    9;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 CTGSAAANYTTAVDRPN 17
          |||    ||||: :| |
Db     14 CTGCKRRNYTTSRNRRN 30


RESULT   8
A60109
major outer membrane protein precursor - Chlamydophila psittaci (strain Guinea pig inclu
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 31-Mar-2000
C;Accession: A60109
R;Zhang, Y.X.; Morrison, S.G.; Caldwell, H.D.; Baehr, W.
Infect. Immun. 57, 1621-1625, 1989
A;Title: Cloning and sequence analysis of the major outer membrane protein genes of two
A;Reference number: A60109; MUID:89212917
A;Accession: A60109
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-389 <ZHA>
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match             51.6%;  Score 47;  DB 2;  Length 389;
  Best Local Similarity   56.2%;  Pred. No. 3.2;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||:|||:: |  || |
Db     89 TGNAAADFKTVADRNN 104


RESULT   9
A84156
single-strand DNA-binding protein (phage-related protein) ssb [imported] - Bacillus h
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000
C;Accession: A84156
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20263314
A;Accession: A84156
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-168 <STO>
A;Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07768.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: ssb
C;Superfamily: bacterial single-stranded DNA-binding protein; single-stranded DNA-bin


  Query Match             50.5%;  Score 46;  DB 2;  Length 168;
  Best Local Similarity   64.3%;  Pred. No. 2;
  Matches    9;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 GSAAANYTTAVDRP 16
          | | ||:| ||:||
Db     23 GVAVANFTLAVNRP 36


RESULT  10
T25753
hypothetical protein F45E4.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25753
R;Wilson, R.
submitted to the EMBL Data Library, September 1996
A;Description: The sequence of C. elegans cosmid F45E4.
A;Reference number: Z20082
A;Accession: T25753
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1325 <WIL>
A;Cross-references: EMBL:U70852; PIDN:AAB09134.1; GSPDB:GN00022; CESP:F45E4.3
A;Experimental source: strain Bristol N2; clone F45E4
C;Genetics:
A;Gene: CESP:F45E4.3
A;Map position: 4
A;Introns: 25/3; 859/1; 928/1; 966/1; 1002/2; 1106/2; 1167/1; 1255/1; 1274/2


  Query Match             50.5%;  Score 46;  DB 2;  Length 1325;
  Best Local Similarity   90.0%;  Pred. No. 17;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 GSAAANYTTA 12
          ||||:|||||
Db    156 GSAASNYTTA 165


RESULT  11
D70575
```

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70575
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: D70575
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-3300 <COL>
A;Cross-references: GB:Z95324; GB:AL123456; NID:g3261760; PIDN:CAB08587.1; PID:e1299834;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: PPE

Query Match 49.5%; Score 45; DB 2; Length 3300;
Best Local Similarity 57.1%; Pred. No. 61;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       3 GSAAANYTTAVDRP 16
           |  |||||  ::||
Db    2792 GLLAANYTITIERP 2805
```

RESULT 12
I39710
cellulose biosynthesis protein celD - Agrobacterium tumefaciens
C;Species: Agrobacterium tumefaciens
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: I39710
R;Matthysse, A.G.; White, S.; Lightfoot, R.
J. Bacteriol. 177, 1069-1075, 1995
A;Title: Genes required for cellulose synthesis in Agrobacterium tumefaciens.
A;Reference number: I39709; MUID:95164506
A;Accession: I39710
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-584 <RES>
A;Cross-references: GB:L38609; NID:g710486; PIDN:AAC41431.1; PID:g710488
C;Comment: This protein is required for cellulose biosynthesis.
C;Genetics:
A;Gene: celD
C;Superfamily: Agrobacterium tumefaciens cellulose biosynthesis protein celD

Query Match 48.4%; Score 44; DB 1; Length 584;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 GSAAANYTTAVDRPN 17
           |:|||:||  : ||:
Db      89 GNAAADYTGFISRPD 103
```

RESULT 13
T01009
hypothetical protein At2g39790 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein T5I7.9
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 23-Mar-2001
C;Accession: T01009; E84821
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, November 1997
A;Description: Arabidopsis thaliana chromosome II BAC T5I7 genomic sequence.
A;Reference number: Z14162
A;Accession: T01009
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-429 <ROU>
A;Cross-references: EMBL:AC003000; NID:g2642152; PID:g2642161
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: E84821
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-429 <STO>
A;Cross-references: GB:AE002093; NID:g2642161; PIDN:AAB87128.1; GSPDB:GN00139
C;Genetics:
A;Gene: T5I7.9; At2g39790
A;Map position: 2
A;Introns: 82/3; 189/2; 263/3; 377/2

Query Match 47.3%; Score 43; DB 2; Length 429;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 SAAANYTTAVDR 15
           | | ||:||:||
Db     226 SRALNYSTAIDR 237
```

RESULT 14
S59540
heat shock transcription factor HSF31 - soybean (fragment)
C;Species: Glycine max (soybean)
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 03-Nov-2000
C;Accession: S59540
R;Czarnecka-Verner, E.; Yuan, C.X.; Fox, P.C.; Gurley, W.B.
Plant Mol. Biol. 29, 37-51, 1995
A;Title: Isolation and characterization of six heat shock transcription factor cDNA c
A;Reference number: S59537; MUID:96017612
A;Accession: S59540
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-306 <CZA>
A;Cross-references: EMBL:Z46955; NID:g662926; PIDN:CAA87079.1; PID:g671867
A;Note: constitutively expressed
C;Genetics:
A;Gene: HSF31
C;Superfamily: tomato heat shock transcription factor HSF8; HSF DNA-binding domain ho
C;Keywords: DNA binding; leucine zipper; nucleus; transcription factor
F;1-63/Domain: HSF DNA-binding domain homology (fragment) <HSF>

Query Match 46.2%; Score 42; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 GSAAANYTTAV 13
           |:|||| |:|
Db     122 GAAAANYNTSV 132
```

RESULT 15
D84403
dihydroorotase [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: D84403
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;

A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483
A;Accession: D84403
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-422 <STO>
A;Cross-references: GB:AE004437; NID:g10581925; PIDN:AAG20592.1; GSPDB:GN00138
C;Genetics:
A;Gene: pyrC
C;Superfamily: Bacillus dihydroorotase; Bacillus dihydroorotase homology

Query Match 46.2%; Score 42; DB 2; Length 422;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       4 SAAANYTTAVDRPN 17
           :||   || ||:||
Db      74 AAAGGVTTVVDQPN 87
```

Search completed: March 26, 2002, 13:37:18
Job time: 52 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:36:26 ; Search time 24.63 Seconds
(without alignments)
25.307 Million cell updates/sec

Title: US-09-709-201-93
Perfect score: 91
Sequence: 1 CTGSAAANYTTAVDRPN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	82	90.1	389	1	OMP1_CHLPN	P27455 chlamydia p
2	78	85.7	333	1	OM1K_CHLPN	Q9xbf4 chlamydia p
3	78	85.7	389	1	OM1N_CHLPN	Q07430 chlamydia p
4	64	70.3	389	1	OM1A_CHLPS	P16567 chlamydia p
5	47	51.6	56	1	RL33_TREPA	O83262 treponema p
6	43	47.3	429	1	YB90_ARATH	O22288 arabidopsis
7	42	46.2	1005	1	Y456_CHLTR	O84462 chlamydia t
8	41	45.1	345	1	PAPA_CARPA	P00784 carica papa
9	41	45.1	477	1	B1AR_HUMAN	P08588 homo sapien
10	41	45.1	480	1	B1AR_MACMU	P47899 macaca mula
11	41	45.1	3969	1	HRX_HUMAN	Q03164 homo sapien
12	40	44.0	87	1	ANSA_STRCZ	P01548 streptomyce
13	40	44.0	280	1	VG27_HSVSA	Q00998 herpesvirus
14	40	44.0	722	1	PALY_CITLI	Q42667 citrus limo
15	39.5	43.4	268	1	DAPB_PSEAE	P38103 pseudomonas
16	39	42.9	375	1	GPRS_HUMAN	Q9ns67 homo sapien
17	39	42.9	377	1	GPRS_RAT	Q9jjh3 rattus norv
18	39	42.9	379	1	GPRS_MOUSE	O54897 mus musculu
19	39	42.9	414	1	Y878_METJA	Q58288 methanococc
20	39	42.9	2037	1	FAS1_CANAL	P34731 c fatty aci
21	38.5	42.3	454	1	PYRC_METTH	O27199 methanobact
22	38	41.8	116	1	Y243_MYCGE	P47485 mycoplasma
23	38	41.8	278	1	YFOL_STRTR	P96051 streptococc
24	38	41.8	309	1	J1L_HCMVA	P17143 human cytom
25	38	41.8	346	1	CRL1_CANAL	P33153 candida alb
26	38	41.8	459	1	ZPR1_SCHPO	O13724 schizosacch
27	38	41.8	538	1	DAC_ACTSP	P39045 actinomadur
28	38	41.8	557	1	YP85_MYCTU	Q50636 mycobacteri
29	38	41.8	580	1	MEND_BACSU	P23970 b menaquino
30	38	41.8	713	1	CDGT_BACSP	P30921 bacillus sp
31	38	41.8	895	1	ANDR_MACFA	O97952 macaca fasc
32	38	41.8	895	1	ANDR_PAPHA	O97960 papio hamad
33	38	41.8	911	1	ANDR_PANTR	O97775 pan troglod
34	38	41.8	919	1	ANDR_HUMAN	P10275 homo sapien
35	38	41.8	1318	1	VP14_EBV	P03179 epstein-bar
36	38	41.8	1790	1	VIT_ANTGR	Q05808 anthonomus
37	37.5	41.2	1142	1	ENAM_PIG	O97939 sus scrofa
38	37	40.7	270	1	NUP1_PENCI	P24289 penicillium
39	37	40.7	270	1	NUP3_PENSQ	P24504 penicillium
40	37	40.7	280	1	YTZ4_AGRVI	P70796 agrobacteri
41	37	40.7	341	1	COA2_POVMK	P24596 mouse polyo
42	37	40.7	505	1	KIR2_HUMAN	P36896 homo sapien
43	37	40.7	718	1	YSO2_CAEEL	Q10128 caenorhabdi
44	37	40.7	754	1	SULX_YEAST	P53394 saccharomyc
45	37	40.7	884	1	CADB_XENLA	P33152 xenopus lae

ALIGNMENTS

RESULT 1
OMP1_CHLPN
ID OMP1_CHLPN STANDARD; PRT; 389 AA.
AC P27455; Q9JQF6;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN OMPA OR OMP1 OR CPN0695 OR CP0051.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IOL-207;
RX MEDLINE=91237311; PubMed=2033374;
RA Carter M.W., Al-Mahdawi S.A.H., Giles I.G., Treharne J.D.,
RA Ward M.E., Clarke I.N.;
RT "Nucleotide sequence and taxonomic value of the major outer membrane
RT protein gene of Chlamydia pneumoniae IOL-207.";
RL J. Gen. Microbiol. 137:465-475(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TWAR;
RX MEDLINE=91244474; PubMed=1840574;
RA Perez Melgosa M., Kuo C.-C., Campbell L.A.;
RT "Sequence analysis of the major outer membrane protein gene of
RT Chlamydia pneumoniae.";
RL Infect. Immun. 59:2195-2199(1991).
RN [3]
RP SEQUENCE FROM N.A.
RA Mitchell W.M., Tharp A.C., Stratton C.W., Sriram S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;

```
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20298986; PubMed=10839753;
RA   Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,
RA   Takeuchi H., Nishida J., Shibata K., Fujinaga R., Yoneda H.,
RA   Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,
RA   Ishii K., Shiba T., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of outer membrane protein genes omp and pmp in the whole
RT   genome sequences of Chlamydia pneumoniae isolates from Japan and the
RT   United States.";
RL   J. Infect. Dis. 181 Suppl 3:S524-S527(2000).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M64064; AAA23143.1; -.
DR   EMBL; M69230; AAA73071.1; -.
DR   EMBL; AF131889; AAD22492.1; -.
DR   EMBL; AE001652; AAD18834.1; -.
DR   EMBL; AE002167; AAF37944.1; -.
DR   EMBL; AP002547; BAA98902.1; -.
DR   EMBL; AB033787; BAA85940.1; -.
DR   PIR; A43587; A43587.
DR   PIR; A49751; A49751.
DR   TIGR; CP0051; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
FT   SIGNAL       1     23
FT   CHAIN       24    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41620 MW;  15D984151E41F8F2 CRC64;


  Query Match             90.1%;  Score 82;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||||||||||||||||
Db     90 TGSAAANYTTAVDRPN 105


RESULT   2
OM1K_CHLPN
ID   OM1K_CHLPN     STANDARD;      PRT;   333 AA.
AC   Q9XBF4;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN (MOMP) (FRAGMENT).
GN   OMPA OR OMP1.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KOALA TYPE I;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M73038; AAD38210.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin.
FT   NON_TER       1       1
FT   NON_TER     333     333
SQ   SEQUENCE   333 AA;  35811 MW;  204604512C4C3B3F CRC64;


  Query Match             85.7%;  Score 78;  DB 1;  Length 333;
  Best Local Similarity   93.8%;  Pred. No. 6.2e-06;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          |||| ||||||||||
Db     46 TGSATANYTTAVDRPN 61


RESULT   3
OM1N_CHLPN
ID   OM1N_CHLPN     STANDARD;      PRT;   389 AA.
AC   Q07430;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N16;
RX   MEDLINE=94103736; PubMed=8277245;
RA   Storey C., Lusher M., Yates P., Richmond S.;
RT   "Evidence for Chlamydia pneumoniae of non-human origin.";
RL   J. Gen. Microbiol. 139:2621-2626(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L04982; AAA17397.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     23       BY SIMILARITY.
FT   CHAIN        24    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41628 MW;  801622F05D841967 CRC64;


  Query Match             85.7%;  Score 78;  DB 1;  Length 389;
  Best Local Similarity   93.8%;  Pred. No. 7.3e-06;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          |||| ||||||||||
Db     90 TGSATANYTTAVDRPN 105


RESULT   4
OM1A_CHLPS
ID   OM1A_CHLPS     STANDARD;      PRT;   389 AA.
AC   P16567;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OVINE ENZOOTIC ABORTION ISOLATE S26/3;
RX   MEDLINE=90128177; PubMed=2612883;
RA   Herring A.J., Tan T.W., Baxter S., Inglis N.F., Dunbar S.;
RT   "Sequence analysis of the major outer membrane protein gene of an
RT   ovine abortion strain of Chlamydia psittaci.";
RL   FEMS Microbiol. Lett. 53:153-158(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BOVINE ABORTION ISOLATE BA1;
RX   MEDLINE=96189695; PubMed=8605581;
RA   Griffiths P.C., Plater J.M., Martin T.C., Hughes S.L.,
RA   Hughes K.J., Hewinson R.G., Dawson M.;
RT   "Epizootic bovine abortion in a dairy herd: characterization of a
RT   Chlamydia psittaci isolate and antibody response.";
RL   Br. Vet. J. 151:683-693(1995).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X51859; CAA36152.1; -.
DR   EMBL; L39020; AAB02850.1; -.
DR   PIR; S08770; MMCWP3.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41883 MW;  741B5A23ACDBB447 CRC64;


  Query Match             70.3%;  Score 64;  DB 1;  Length 389;
  Best Local Similarity   75.0%;  Pred. No. 0.0017;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||:||||| |  ||||
Db     90 TGTAAANYKTPTDRPN 105


RESULT   5
RL33_TREPA
ID   RL33_TREPA     STANDARD;      PRT;    56 AA.
AC   O83262;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   50S RIBOSOMAL PROTEIN L33.
GN   RPMG OR TP0234.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NICHOLS;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- SIMILARITY: BELONGS TO THE L33P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE001205; AAC65222.1; -.
DR   TIGR; TP0234; -.
DR   InterPro; IPR001705; Ribosomal_L33.
DR   Pfam; PF00471; Ribosomal_L33; 1.
DR   ProDom; PD002595; Ribosomal_L33; 1.
DR   PROSITE; PS00582; RIBOSOMAL_L33; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   56 AA;  6820 MW;  1636DC3500D1F4B1 CRC64;


  Query Match             51.6%;  Score 47;  DB 1;  Length 56;
  Best Local Similarity   52.9%;  Pred. No. 0.2;
  Matches    9;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 CTGSAAANYTTAVDRPN 17
          |||    ||||: :| |
Db     14 CTGCKRRNYTTSRNRRN 30
```

```
RESULT   6
YB90_ARATH
ID   YB90_ARATH     STANDARD;      PRT;   429 AA.
AC   O22288;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL 47.9 KDA PROTEIN AT2G39790.
GN   AT2G39790 OR T5I7.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., VanAken S.E., Umayam L.,
RA   Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
CC   -!- SIMILARITY: DISTANTLY RELATED TO THE MAM33 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AC003000; AAB87128.1; -.
DR   InterPro; IPR003428; MAM33.
DR   Pfam; PF02330; MAM33; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   429 AA;  47860 MW;  FC5F0BD011CF0A83 CRC64;


  Query Match             47.3%;  Score 43;  DB 1;  Length 429;
  Best Local Similarity   66.7%;  Pred. No. 7;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 SAAANYTTAVDR 15
           | | ||:||:||
Db     226 SRALNYSTAIDR 237


RESULT   7
Y456_CHLTR
ID   Y456_CHLTR     STANDARD;      PRT;  1005 AA.
AC   O84462;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL PROTEIN CT456 PRECURSOR.
GN   CT456.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/CX;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CT456/TC0741
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001319; AAC68056.1; -.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41   1005       HYPOTHETICAL PROTEIN CT456.
SQ   SEQUENCE   1005 AA;  102131 MW;  EC47EC389851CD1E CRC64;


  Query Match             46.2%;  Score 42;  DB 1;  Length 1005;
  Best Local Similarity   43.8%;  Pred. No. 24;
  Matches    7;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
             ||    ||| :   ||
Db      60 TGETVVNYTNSASAPN 75


RESULT   8
PAPA_CARPA
ID   PAPA_CARPA     STANDARD;      PRT;   345 AA.
AC   P00784;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PAPAIN PRECURSOR (EC 3.4.22.2) (PAPAYA PROTEINASE I) (PPI).
OS   Carica papaya (Papaya).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Caricaceae; Carica.
OX   NCBI_TaxID=3649;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87163525; PubMed=2881845;
RA   Cohen L.W., Coghlan V.M., Dihel L.C.;
RT   "Cloning and sequencing of papain-encoding cDNA.";
RL   Gene 48:219-227(1986).
RN   [2]
RP   SEQUENCE OF 134-345.
RX   MEDLINE=71007899; PubMed=5470818;
RA   Mitchel R.E.J., Chaiken I.M., Smith E.L.;
RT   "The complete amino acid sequence of papain. Additions and
RT   corrections.";
RL   J. Biol. Chem. 245:3485-3492(1970).
RN   [3]
RP   REVISION TO 197.
RX   MEDLINE=70141125; PubMed=5435495;
RA   Husain S.S., Lowe G.;
RT   "A reinvestigation of residues 64-68 and 175 in papain. Evidence that
RT   residues 64 and 175 are asparagine.";
RL   Biochem. J. 116:689-692(1970).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX   MEDLINE=69006973; PubMed=5681232;
RA   Drenth J., Jansonius J.N., Koekoek R., Swen H.M., Wolthers B.G.;
RT   "Structure of papain.";
RL   Nature 218:929-932(1968).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RX   MEDLINE=85058190; PubMed=6502713;
```

```
RA   Kamphuis I.G., Kalk K.H., Swarte M.B.A., Drenth J.;
RT   "Structure of papain refined at 1.65-A resolution.";
RL   J. Mol. Biol. 179:233-256(1984).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX   MEDLINE=90269230; PubMed=2347312;
RA   Stubbs M.T., Laber B., Bode W., Huber R., Jerala R., Lenarcic B.,
RA   Turk V.;
RT   "The refined 2.4 A X-ray crystal structure of recombinant human
RT   stefin B in complex with the cysteine proteinase papain: a novel type
RT   of proteinase inhibitor interaction.";
RL   EMBO J. 9:1939-1947(1990).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX   MEDLINE=93075728; PubMed=1445868;
RA   Yamamoto A., Tomoo K., Doi M., Ohishi H., Inoue M., Ishida T.,
RA   Yamamoto D., Tsuboi S., Okamoto H., Okada Y.;
RT   "Crystal structure of
RT   papain-succinyl-Gln-Val-Val-Ala-Ala-p-nitroanilide complex at 1.7-A
RT   resolution: noncovalent binding mode of a common sequence of
RT   endogenous thiol protease inhibitors.";
RL   Biochemistry 31:11305-11309(1992).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA   Pickersgill R.W., Harris G.W., Garman E.;
RT   "Structure of monoclinic papain at 1.60-A resolution.";
RL   Acta Crystallogr. B 48:59-67(1992).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-. LYS-, PHE-XAA-.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC       PAPAIN FAMILY OF THIOL PROTEASES.
CC   -!- DATABASE: NAME=Worthington enzyme manual;
CC       WWW="http://www.worthington-biochem.com/manual/P/PAP.html".
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M15203; AAB02650.1; -.
DR   PIR; A26466; PPPA.
DR   PDB; 1PAD; 17-FEB-84.
DR   PDB; 2PAD; 17-FEB-84.
DR   PDB; 4PAD; 04-MAR-85.
DR   PDB; 5PAD; 17-FEB-84.
DR   PDB; 6PAD; 17-FEB-84.
DR   PDB; 9PAP; 15-JAN-95.
DR   PDB; 1PE6; 15-APR-93.
DR   PDB; 1PIP; 31-OCT-93.
DR   PDB; 1POP; 31-OCT-93.
DR   PDB; 1PPD; 02-JAN-85.
DR   PDB; 1PPN; 31-JAN-94.
DR   PDB; 1PPP; 31-JAN-94.
DR   PDB; 1STF; 31-JAN-94.
DR   MEROPS; C01.001; -.
DR   InterPro; IPR000668; Peptidase_C1.
DR   InterPro; IPR000169; Thiolprot_act_site.
DR   Pfam; PF00112; Peptidase_C1; 1.
DR   PRINTS; PR00705; PAPAIN.
DR   PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR   PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR   PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
KW   Hydrolase; Thiol protease; Zymogen; Signal; 3D-structure.
FT   SIGNAL       1     18       POTENTIAL.
FT   PROPEP      19    133       ACTIVATION PEPTIDE.
FT   CHAIN      134    345       PAPAIN.
FT   ACT_SITE   158    158
FT   ACT_SITE   292    292
FT   ACT_SITE   308    308
FT   DISULFID   155    196
FT   DISULFID   189    228
FT   DISULFID   286    333
FT   CONFLICT   180    180       E -> Q (IN REF. 2).
FT   CONFLICT   219    220       YP -> PY (IN REF. 2).
FT   CONFLICT   251    251       E -> Q (IN REF. 2).
FT   CONFLICT   268    268       E -> Q (IN REF. 2).
FT   STRAND     138    139
FT   TURN       140    144
FT   STRAND     151    151
FT   TURN       153    154
FT   STRAND     156    156
FT   HELIX      158    175
FT   STRAND     181    181
FT   HELIX      183    189
FT   TURN       191    192
FT   TURN       195    196
FT   STRAND     197    197
FT   STRAND     199    199
FT   HELIX      201    210
FT   TURN       211    211
FT   STRAND     213    213
FT   STRAND     215    215
FT   TURN       216    218
FT   HELIX      231    233
FT   STRAND     238    238
FT   STRAND     242    245
FT   HELIX      251    260
FT   STRAND     263    267
FT   HELIX      272    276
FT   STRAND     281    282
FT   STRAND     292    299
FT   STRAND     303    307
FT   STRAND     310    310
FT   TURN       312    313
FT   TURN       315    315
FT   STRAND     316    316
FT   TURN       317    318
FT   STRAND     319    323
FT   HELIX      332    334
FT   TURN       335    336
FT   STRAND     340    343
SQ   SEQUENCE   345 AA;  38922 MW;  82D9FB35EDCA12EF CRC64;


  Query Match           45.1%;  Score 41;  DB 1;  Length 345;
  Best Local Similarity  80.0%;  Pred. No. 12;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTT 11
           ||| | ::|:|
Db     109 TGSIAGNYTT 118


RESULT   9
B1AR_HUMAN
ID   B1AR_HUMAN     STANDARD;      PRT;   477 AA.
AC   P08588; O9UKG8; O9UKG7;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   BETA-1 ADRENERGIC RECEPTOR.
GN   ADRB1 OR ADRB1R OR B1AR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=88068509; PubMed=2825170;
RA   Frielle T., Collins S., Daniel K.W., Caron M.G., Lefkowitz R.J.,
RA   Kobilka B.K.;
RT   "Cloning of the cDNA for the human beta 1-adrenergic receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:7920-7924(1987).
```

```
RN   [2]
RP   VARIANT ARG-389.
RX   MEDLINE=99230291; PubMed=10212248;
RA   Mason D.A., Moore J.D., Green S.A., Liggett S.B.;
RT   "A gain-of-function polymorphism in a G-protein coupling domain of the
RT   human beta1-adrenergic receptor.";
RL   J. Biol. Chem. 274:12670-12674(1999).
RN   [3]
RP   VARIANTS GLY-49 AND ARG-389.
RX   MEDLINE=99407229; PubMed=10477438;
RA   Moore J.D., Mason D.A., Green S.A., Hsu J., Liggett S.B.;
RT   "Racial differences in the frequencies of cardiac beta(1)-adrenergic
RT   receptor polymorphisms: analysis of c145A>G and c1165G>C.";
RL   Hum. Mutat. 14:271-271(1999).
RN   [4]
RP   VARIANT GLY-49.
RX   MEDLINE=20507547; PubMed=11052857;
RA   Borjesson M., Magnusson Y., Hjalmarson A., Andersson B.;
RT   "A novel polymorphism in the gene coding for the beta(1)-adrenergic
RT   receptor associated with survival in patients with heart failure.";
RL   Eur. Heart J. 21:1853-1858(2000).
CC   -!- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC       INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC       PROTEINS. THIS RECEPTOR BINDS EPINEPHRINE AND NOREPINEPHRINE WITH
CC       APPROXIMATIVELY EQUAL AFFINITY.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- PTM: HOMOLOGOUS DESENSITIZATION OF THE RECEPTOR IS MEDIATED BY ITS
CC       PHOSPHORYLATION BY BETA-ADRENERGIC RECEPTOR KINASE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03019; AAA51667.1; -.
DR   EMBL; AF169006; AAD53696.1; -.
DR   EMBL; AF169007; AAD53697.1; -.
DR   PIR; A39911; A39911.
DR   HSSP; P07700; 1DEP.
DR   GCRDb; GCR_0048; -.
DR   MIM; 109630; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00561; ADRENRGCB1AR.
DR   PRINTS; PR01103; ADRENERGICR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Phosphorylation; Lipoprotein; Palmitate;
KW   Polymorphism.
FT   DOMAIN       1     59       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     60    83       1 (POTENTIAL).
FT   DOMAIN       84    96       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     97    120      2 (POTENTIAL).
FT   DOMAIN       121   131      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     132   155      3 (POTENTIAL).
FT   DOMAIN       156   175      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     176   199      4 (POTENTIAL).
FT   DOMAIN       200   221      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     222   245      5 (POTENTIAL).
FT   DOMAIN       246   325      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     326   349      6 (POTENTIAL).
FT   DOMAIN       350   356      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     357   380      7 (POTENTIAL).
FT   DOMAIN       381   477      CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      15    15      N-LINKED (GLCNAC. . .) (PROBABLE).
FT   DISULFID     131   209      BY SIMILARITY.
FT   MOD_RES      312   312      PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT   MOD_RES      412   412      PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT   LIPID        392   392      PALMITATE (BY SIMILARITY).
FT   VARIANT       49    49      S -> G.
FT                               /FTId=VAR_009879.
FT   VARIANT      389   389      G -> R (ENHANCED BINDING TO G PROTEINS).
FT                               /FTId=VAR_009880.
SQ   SEQUENCE   477 AA;  51223 MW;  1D15E6390B5364B8 CRC64;


  Query Match             45.1%;  Score 41;  DB 1;  Length 477;
  Best Local Similarity   43.8%;  Pred. No. 17;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDRP 16
           | | |||:  :::| |
Db     451 CNGGAAADSDSSLDEP 466


RESULT  10
B1AR_MACMU
ID   B1AR_MACMU      STANDARD;      PRT;   480 AA.
AC   P47899;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   BETA-1 ADRENERGIC RECEPTOR.
GN   ADRB1.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95078456; PubMed=7987008;
RA   Searles R.P., Nipper V.J., Machida C.A.;
RT   "The rhesus macaque beta 1-adrenergic receptor gene: structure of the
RT   gene and comparison of the flanking sequences with the rat beta 1-
RT   adrenergic receptor gene.";
RL   DNA Seq. 4:231-241(1994).
CC   -!- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC       INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC       PROTEINS. THIS RECEPTOR BINDS EPINEPHRINE AND NOREPINEPHRINE WITH
CC       APPROXIMATIVELY EQUAL AFFINITY.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- PTM: HOMOLOGOUS DESENSITIZATION OF THE RECEPTOR IS MEDIATED BY ITS
CC       PHOSPHORYLATION BY BETA-ADRENERGIC RECEPTOR KINASE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75540; CAA53228.1; -.
DR   HSSP; P07700; 1DEP.
DR   GCRDb; GCR_1746; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00561; ADRENRGCB1AR.
DR   PRINTS; PR01103; ADRENERGICR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN       1     59       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     60    83       1 (POTENTIAL).
FT   DOMAIN       84    96       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     97    120      2 (POTENTIAL).
```

```
FT   DOMAIN      121    131       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    132    155       3 (POTENTIAL).
FT   DOMAIN      156    175       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    176    199       4 (POTENTIAL).
FT   DOMAIN      200    221       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    222    245       5 (POTENTIAL).
FT   DOMAIN      246    328       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    329    352       6 (POTENTIAL).
FT   DOMAIN      353    359       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    360    383       7 (POTENTIAL).
FT   DOMAIN      384    480       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     15     15       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   DISULFID    131    209       BY SIMILARITY.
FT   MOD_RES     315    315       PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT   MOD_RES     415    415       PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT   LIPID       395    395       PALMITATE (BY SIMILARITY).
SQ   SEQUENCE   480 AA;  51608 MW;  25CB18FA03128084 CRC64;


  Query Match             45.1%;  Score 41;  DB 1;  Length 480;
  Best Local Similarity   43.8%;  Pred. No. 17;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVDRP 16
           | | |||:  :::| |
Db     454 CNGGAAADSDSSLDEP 469


RESULT  11
HRX_HUMAN
ID   HRX_HUMAN      STANDARD;      PRT;  3969 AA.
AC   Q03164; Q14845; Q16364; Q13743; Q13744; Q9UMA3;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   ZINC FINGER PROTEIN HRX (ALL-1) (TRITHORAX-LIKE PROTEIN).
GN   MLL OR HRX OR ALL1 OR TRX1 OR HTRX.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93046667; PubMed=1423624;
RA   Tkachuk D.C., Kohler S., Cleary M.L.;
RT   "Involvement of a homolog of Drosophila trithorax by 11q23
RT   chromosomal translocations in acute leukemias.";
RL   Cell 71:691-700(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96290553; PubMed=8703835;
RA   Nilson I., Loechner K., Siegler G., Greil J., Beck J.D., Fey G.H.,
RA   Marschalek R.;
RT   "Exon/intron structure of the human ALL-1 (MLL) gene involved in
RT   translocations to chromosomal region 11q23 and acute leukaemias.";
RL   Br. J. Haematol. 93:966-972(1996).
RN   [3]
RP   SEQUENCE OF 1-1909 FROM N.A.
RX   MEDLINE=93390935; PubMed=8378076;
RA   Yamamoto K., Seto M., Komatsu H., Iida S., Akao Y., Kojima S.,
RA   Kodera Y., Nakazawa S., Ariyoshi Y., Takahashi T., Ueda R.;
RT   "Two distinct portions of LTG19/ENL at 19p13 are involved in t(11;19)
RT   leukemia.";
RL   Oncogene 8:2617-2625(1993).
RN   [4]
RP   SEQUENCE OF 1317-2328 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=93265134; PubMed=1303259;
RA   Djabali M., Selleri L., Parry P., Bower M., Young B.D., Evans G.A.;
RT   "A trithorax-like gene is interrupted by chromosome 11q23
RT   translocations in acute leukaemias.";
RL   Nat. Genet. 2:113-118(1992).
RN   [5]
RP   SEQUENCE OF 1251-1538 FROM N.A.
RX   MEDLINE=94215165; PubMed=8162575;
RA   Gu Y., Alder H., Nakamura T., Schichman S.A., Prasad R., Canaani O.,
RA   Saito H., Croce C.M., Canaani E.;
RT   "Sequence analysis of the breakpoint cluster region in the ALL-1 gene
RT   involved in acute leukemia.";
RL   Cancer Res. 54:2326-2330(1994).
RN   [6]
RP   SEQUENCE OF 1251-1654 FROM N.A. (ISOFORM 14P-18B).
RX   MEDLINE=95322025; PubMed=7598802;
RA   Mbangkollo D., Burnett R., McCabe N., Thirman M., Gill H., Yu H.,
RA   Rowley J.D., Diaz M.O.;
RT   "The human MLL gene: nucleotide sequence, homology to the Drosophila
RT   trx zinc-finger domain, and alternative splicing.";
RL   DNA Cell Biol. 14:475-483(1995).
RN   [7]
RP   SEQUENCE OF 1212-1603 FROM N.A.
RX   MEDLINE=95315013; PubMed=7794749;
RA   Marschalek R., Greil J., Lochner K., Nilson I., Siegler G.,
RA   Zweckbronner I., Beck J.D., Fey G.H.;
RT   "Molecular analysis of the chromosomal breakpoint and fusion
RT   transcripts in the acute lymphoblastic SEM cell line with chromosomal
RT   translocation t(4;11).";
RL   Br. J. Haematol. 90:308-320(1995).
RN   [8]
RP   SEQUENCE OF 1421-1540 FROM N.A.
RX   MEDLINE=94020842; PubMed=8414518;
RA   Forster A., Rabbitts T.H.;
RT   "A method for identifying genes within yeast artificial chromosomes:
RT   application to isolation of MLL fusion cDNAs from acute leukaemia
RT   translocations.";
RL   Oncogene 8:3157-3160(1993).
RN   [9]
RP   CHROMOSOMAL TRANSLOCATION WITH GAS7.
RX   MEDLINE=20183971; PubMed=10706619;
RA   Megonigal M.D., Cheung N.-K.V., Rappaport E.F., Nowell P.C.,
RA   Wilson R.B., Jones D.H., Addya K., Leonard D.G.B., Kushner B.H.,
RA   Williams T.M., Lange B.J., Felix C.A.;
RT   "Detection of leukemia-associated MLL-GAS7 translocation early during
RT   chemotherapy with DNA topoisomerase II inhibitors.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2814-2819(2000).
CC   -!- FUNCTION: POSSIBLY ACTS AS A TRANSCRIPTIONAL REGULATORY FACTOR.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: HEART, LUNG, BRAIN AND T AND B LYMPHOCYTES.
CC   -!- DISEASE: INVOLVED IN ACUTE LEUKEMIAS BY CHROMOSOMAL TRANSLOCATIONS
CC       T(11;19)(Q23;P13.3) THAT INVOLVES MLL AND MLLT1/ENL;
CC       T(4;11)(Q21;Q23) THAT INVOLVES MLL AND MLLT2/AF4; T(9;11)(P22;Q23)
CC       THAT INVOLVES MLL AND MLLT3/AF9; T(6;11)(Q27;Q23) THAT INVOLVES
CC       MLL AND MLLT4/AF6; T(11;17)(Q23;Q21) THAT INVOLVES MLL AND
CC       MLLT6/AF17; T(X;11)(Q13;Q23) THAT INVOLVES MLL AND MLLT7/AFX1;
CC       T(10;11)(P12;Q23) THAT INVOLVES MLL AND MLLT10/AF10;
CC       T(1;11)(Q21;Q23) THAT INVOLVES MLL AND AF1Q; T(11;19)(Q23;P13.3)
CC       THAT INVOLVES MLL AND ELL; AND T(11;19)(Q23;P23) THAT INVOLVES MLL
CC       AND GAS7.
CC   -!- SIMILARITY: BELONGS TO THE TRITHORAX FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 BROMODOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SET DOMAIN.
CC   -!- SIMILARITY: CONTAINS 3 PHD-TYPE ZINC FINGERS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L04284; AAA58669.1; -.
DR   EMBL; Z69744; CAA93625.1; -.
DR   EMBL; Z69745; CAA93625.1; JOINED.
DR   EMBL; Z69746; CAA93625.1; JOINED.
DR   EMBL; Z69747; CAA93625.1; JOINED.
```

```
DR   EMBL; Z69748; CAA93625.1; JOINED.
DR   EMBL; Z69749; CAA93625.1; JOINED.
DR   EMBL; Z69750; CAA93625.1; JOINED.
DR   EMBL; Z69751; CAA93625.1; JOINED.
DR   EMBL; Z69752; CAA93625.1; JOINED.
DR   EMBL; Z69753; CAA93625.1; JOINED.
DR   EMBL; Z69754; CAA93625.1; JOINED.
DR   EMBL; Z69755; CAA93625.1; JOINED.
DR   EMBL; Z69756; CAA93625.1; JOINED.
DR   EMBL; Z69757; CAA93625.1; JOINED.
DR   EMBL; Z69758; CAA93625.1; JOINED.
DR   EMBL; Z69759; CAA93625.1; JOINED.
DR   EMBL; Z69760; CAA93625.1; JOINED.
DR   EMBL; Z69761; CAA93625.1; JOINED.
DR   EMBL; Z69762; CAA93625.1; JOINED.
DR   EMBL; Z69763; CAA93625.1; JOINED.
DR   EMBL; Z69764; CAA93625.1; JOINED.
DR   EMBL; Z69765; CAA93625.1; JOINED.
DR   EMBL; Z69766; CAA93625.1; JOINED.
DR   EMBL; Z69767; CAA93625.1; JOINED.
DR   EMBL; Z69768; CAA93625.1; JOINED.
DR   EMBL; Z69769; CAA93625.1; JOINED.
DR   EMBL; Z69770; CAA93625.1; JOINED.
DR   EMBL; Z69772; CAA93625.1; JOINED.
DR   EMBL; Z69773; CAA93625.1; JOINED.
DR   EMBL; Z69774; CAA93625.1; JOINED.
DR   EMBL; Z69775; CAA93625.1; JOINED.
DR   EMBL; Z69776; CAA93625.1; JOINED.
DR   EMBL; Z69777; CAA93625.1; JOINED.
DR   EMBL; Z69778; CAA93625.1; JOINED.
DR   EMBL; Z69779; CAA93625.1; JOINED.
DR   EMBL; Z69780; CAA93625.1; JOINED.
DR   EMBL; D14540; BAA03407.1; -.
DR   EMBL; L01986; AAA92511.1; -.
DR   EMBL; U04737; AAA18644.1; -.
DR   EMBL; S78570; AAB34770.1; -.
DR   EMBL; X83604; CAA58584.1; -.
DR   EMBL; S66432; AAB28545.1; -.
DR   EMBL; AF231998; AAG26332.2; ALT_TERM.
DR   TRANSFAC; T02337; -.
DR   MIM; 159555; -.
DR   InterPro; IPR001487; Bromodomain.
DR   InterPro; IPR003889; FYrich_C.
DR   InterPro; IPR003888; FYrich_N.
DR   InterPro; IPR001965; PHD.
DR   InterPro; IPR003616; PostSET.
DR   InterPro; IPR001214; SET.
DR   InterPro; IPR002857; Znf-CXXC.
DR   Pfam; PF00628; PHD; 3.
DR   Pfam; PF00856; SET; 1.
DR   Pfam; PF02008; zf-CXXC; 1.
DR   SMART; SM00297; BROMO; 1.
DR   SMART; SM00542; FYRC; 1.
DR   SMART; SM00541; FYRN; 1.
DR   SMART; SM00249; PHD; 4.
DR   SMART; SM00508; PostSET; 1.
DR   SMART; SM00317; SET; 1.
DR   PROSITE; PS50014; BROMODOMAIN_2; 1.
DR   PROSITE; PS50280; SET; 1.
KW   Proto-oncogene; Chromosomal translocation; DNA-binding; Bromodomain;
KW   Nuclear protein; Zinc-finger; Metal-binding; Transcription regulation;
KW   Alternative splicing.
FT   DOMAIN       17     102       ALA/GLY/SER-RICH.
FT   DNA_BIND     169    180       A.T HOOK (BY SIMILARITY).
FT   DNA_BIND     217    227       A.T HOOK (BY SIMILARITY).
FT   DNA_BIND     301    309       A.T HOOK (BY SIMILARITY).
FT   ZN_FING     1431   1482       PHD-TYPE 1.
FT   ZN_FING     1484   1533       PHD-TYPE 2.
FT   ZN_FING     1566   1627       PHD-TYPE 3.
FT   DOMAIN      1703   1748       BROMODOMAIN (DIVERGENT).
FT   DOMAIN      3840   3969       SET.
FT   DOMAIN       137    143       POLY-GLY.
FT   DOMAIN       561    564       POLY-PRO.
FT   DOMAIN       568    571       POLY-PRO.
FT   SITE        1444   1445       BREAKPOINT FOR TRANSLOCATION TO FORM MLL-
FT                                 GAS7 ONCOGENE.
FT   VARSPLIC    1407   1444       MISSING (IN ISOFORM 14P-18B).
FT   CONFLICT     144    144       E -> ELTTQIPCSWRTKGHIHDKKTEPFRLLAWSWCLN
FT                                 (IN REF. 2).
FT   CONFLICT     317    379       GLLINSELEKPQKVRKDKEGTPPLTKEDKTVVRQSPRRIKP
FT                                 VRIIPSSKRTDATIAKQLLQRA -> VSSLILNWKSPRKSG
FT                                 KTRKEHLHLQKKIRQLSDKALEGLSQLGLFLLQKGQMQPLL
FT                                 SNSYRGQ (IN REF. 1).
FT   CONFLICT     556    556       Q -> E (IN REF. 2).

  Query Match             45.1%;  Score 41;  DB 1;  Length 3969;
  Best Local Similarity   43.8%;  Pred. No. 1.4e+02;
  Matches    7;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           |   :|:|  || ||
Db    3476 SGPQVSNFTQTVDAPN 3491


RESULT  12
ANSA_STRCZ
ID   ANSA_STRCZ      STANDARD;      PRT;    87 AA.
AC   P01548;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   20-MAR-1987 (Rel. 04, Last annotation update)
DE   ANTIBACTERIAL SUBSTANCE A.
OS   Streptomyces carzinostaticus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1897;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=F41;
RX   MEDLINE=70039118; PubMed=5353565;
RA   Sato H., Tanimura T., Nakajima T., Tamura Z.;
RT   "The total amino acid sequence of substance A produced by
RT   Streptomyces carzinostaticus.";
RL   Chem. Pharm. Bull. 17:2188-2191(1969).
DR   PIR; A01811; BXSMAC.
KW   Antibiotic.
FT   DISULFID     42     50
SQ   SEQUENCE   87 AA;  8477 MW;  C9A114BE1534029B CRC64;


  Query Match             44.0%;  Score 40;  DB 1;  Length 87;
  Best Local Similarity   66.7%;  Pred. No. 4.7;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAV 13
           || | | |:|||
Db       8 TGGAVATYSTAV 19


RESULT  13
VG27_HSVSA
ID   VG27_HSVSA      STANDARD;      PRT;    280 AA.
AC   Q00998;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   HYPOTHETICAL GENE 27 PROTEIN.
GN   27.
OS   Herpesvirus saimiri (strain 11).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=10383;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92333688; PubMed=1321287;
RA   Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
```

```
RA   Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA   Honess R.W.;
RT   "Primary structure of the herpesvirus saimiri genome.";
RL   J. Virol. 66:5047-5058(1992).
CC   -!- SIMILARITY: LOW, TO EBV BDLF2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X64346; CAA45650.1; -.
DR   PIR; G36808; G36808.
KW   Hypothetical protein.
SQ   SEQUENCE   280 AA;  32372 MW;  6B470950E93E9ABF CRC64;


  Query Match             44.0%;  Score 40;  DB 1;  Length 280;
  Best Local Similarity   50.0%;  Pred. No. 15;
  Matches    7;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDR 15
           || ::|||  |::|
Db     146 TGGSSANYKLALER 159


RESULT  14
PALY_CITLI
ID   PALY_CITLI     STANDARD;      PRT;   722 AA.
AC   Q42667;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PHENYLALANINE AMMONIA-LYASE (EC 4.3.1.5).
GN   PAL6.
OS   Citrus limon (Lemon).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Sapindales; Rutaceae; Citrus.
OX   NCBI_TaxID=2708;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Seelenfreund D., Chiong M., Lobos S., Perez L.M.;
RT   "A full-length cDNA coding for phenylalanine ammonia-lyase from Citrus
RT   limon.";
RL   (In) Plant Gene Register PGR96-026.
CC   -!- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC       FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC       VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC   -!- CATALYTIC ACTIVITY: L-PHENYLALANINE = TRANS-CINNAMATE + NH(3).
CC   -!- PATHWAY: KEY ENZYME OF PHENYLPROPANOID METABOLISM.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC   -!- PTM: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
CC       WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC       OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U43338; AAB67733.1; -.
DR   InterPro; IPR001106; PAL.
DR   Pfam; PF00221; PAL; 1.
DR   PROSITE; PS00488; PAL_HISTIDASE; 1.
KW   Lyase; Phenylpropanoid metabolism; Multigene family.
FT   SITE        206    208       MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-
FT                                5-ONE (BY SIMILARITY).
SQ   SEQUENCE   722 AA;  78490 MW;  C96893196E30D9E5 CRC64;


  Query Match             44.0%;  Score 40;  DB 1;  Length 722;
  Best Local Similarity   50.0%;  Pred. No. 38;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CTGSAAANYTTAVD 14
           |||:   |:| | |
Db      25 CTGTDPLNWTVAAD 38


RESULT  15
DAPB_PSEAE
ID   DAPB_PSEAE     STANDARD;      PRT;   268 AA.
AC   P38103;
DT   01-OCT-1994 (Rel. 30, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   DIHYDRODIPICOLINATE REDUCTASE (EC 1.3.1.26) (DHPR).
GN   DAPB OR PA4759.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
RN   [2]
RP   SEQUENCE OF 135-268 FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=94222830; PubMed=8169201;
RA   Kwon D.-H., Lu C.-D., Walthall D.A., Brown T.M., Houghton J.E.,
RA   Abdelal A.T.;
RT   "Structure and regulation of the carAB operon in Pseudomonas
RT   aeruginosa and Pseudomonas stutzeri: no untranslated region exists.";
RL   J. Bacteriol. 176:2532-2542(1994).
CC   -!- CATALYTIC ACTIVITY: 2,3,4,5-TETRAHYDRODIPICOLINATE + NAD(P)(+) =
CC       2,3-DIHYDRODIPICOLINATE + NAD(P)H.
CC   -!- PATHWAY: BIOSYNTHESIS OF DIAMINOPIMELATE AND LYSINE FROM ASPARTATE
CC       SEMIALDEHYDE; SECOND STEP.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE DIHYDRODIPICOLINATE REDUCTASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE004889; AAG08145.1; -.
DR   EMBL; U04992; AAA19045.1; -.
DR   EMBL; U81259; AAB39249.1; -.
DR   HSSP; P04036; 1DRV.
DR   InterPro; IPR000846; DapB.
DR   Pfam; PF01113; DapB; 1.
DR   ProDom; PD004105; DapB; 1.
DR   PROSITE; PS01298; DAPB; 1.
KW   Diaminopimelate biosynthesis; Lysine biosynthesis; Oxidoreductase;
```

KW NADP; Complete proteome.
SQ SEQUENCE 268 AA; 28324 MW; 0B37EAF688419254 CRC64;

Query Match 43.4%; Score 39.5; DB 1; Length 268;
Best Local Similarity 62.5%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

```
Qy       2 TGSAAANYTTAVDRPN 17
           ||  ||  | |||||:
Db      24 TG-GAAGLTAAVDRPD 38
```

Search completed: March 26, 2002, 13:40:42
Job time: 256 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:36:26 ; Search time 79.01 Seconds
(without alignments)
31.472 Million cell updates/sec

Title: US-09-709-201-93
Perfect score: 91
Sequence: 1 CTGSAAANYTTAVDRPN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	78	85.7	389	2	Q08085	Q08085 chlamydia p
2	64	70.3	341	2	Q9X717	Q9x717 chlamydophi
3	64	70.3	352	2	O69306	O69306 chlamydia p
4	64	70.3	352	2	O69307	O69307 chlamydia p
5	64	70.3	352	2	O70050	O70050 chlamydia p
6	64	70.3	352	2	O70085	O70085 chlamydia p
7	64	70.3	353	2	O69305	O69305 chlamydia p
8	62	68.1	388	2	Q9AIK1	Q9aik1 chlamydia p
9	59	64.8	389	2	Q9APM4	Q9apm4 chlamydophi
10	48	52.7	557	2	Q99V50	Q99v50 staphylococ
11	48	52.7	662	4	Q9UFR7	Q9ufr7 homo sapien
12	48	52.7	1241	4	Q14148	Q14148 homo sapien
13	47	51.6	326	2	Q9K5C5	Q9k5c5 chlamydia p
14	47	51.6	389	2	Q9AIH9	Q9aih9 chlamydophi
15	46	50.5	168	2	Q9K5N9	Q9k5n9 bacillus ha
16	46	50.5	836	5	Q9BI65	Q9bi65 caenorhabdi
17	46	50.5	1231	5	Q9BI66	Q9bi66 caenorhabdi
18	45	49.5	3300	2	O06304	O06304 mycobacteri
19	44.5	48.9	356	2	O52924	O52924 chlamydia p
20	44.5	48.9	390	2	Q9AIJ5	Q9aij5 chlamydia p
21	44.5	48.9	392	2	Q9AIJ4	Q9aij4 chlamydia p
22	44	48.4	584	2	Q59152	Q59152 agrobacteri
23	43	47.3	1459	3	Q9HG03	Q9hg03 penicillium
24	42	46.2	306	10	Q43456	Q43456 glycine max
25	42	46.2	388	2	Q9AIK0	Q9aik0 chlamydia p
26	42	46.2	391	2	Q46235	Q46235 chlamydia p
27	42	46.2	422	1	Q9HMH9	Q9hmh9 halobacteri
28	42	46.2	499	5	Q9VWW1	Q9vww1 drosophila
29	42	46.2	618	2	Q9A9F7	Q9a9f7 caulobacter
30	42	46.2	1583	12	O90304	O90304 breda virus
31	41	45.1	218	2	O34189	O34189 neisseria g
32	41	45.1	229	2	P96873	P96873 mycobacteri
33	41	45.1	273	2	Q9JWH6	Q9jwh6 neisseria m
34	41	45.1	310	3	Q9P356	Q9p356 lentinula e
35	41	45.1	385	12	O39888	O39888 hepatitis b
36	41	45.1	423	2	Q9JXG1	Q9jxg1 neisseria m
37	41	45.1	431	2	O85283	O85283 sphingomona
38	41	45.1	589	3	Q9HE13	Q9he13 schizosacch
39	41	45.1	626	2	Q45877	Q45877 clostridium
40	41	45.1	693	13	Q91889	Q91889 xenopus lae
41	41	45.1	834	3	Q9P978	Q9p978 neurospora
42	41	45.1	1682	5	Q9V693	Q9v693 drosophila
43	41	45.1	3010	12	Q9J3G2	Q9j3g2 hepatitis c
44	40	44.0	337	1	Q9HP24	Q9hp24 halobacteri
45	40	44.0	375	2	Q9HXE3	Q9hxe3 pseudomonas

ALIGNMENTS

RESULT 1
Q08085
ID Q08085 PRELIMINARY; PRT; 389 AA.
AC Q08085;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
OS Chlamydia psittaci (Chlamydophila psittaci).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KOALA TYPE 1;
RX MEDLINE=94171025; PubMed=8125292;
RA Girjes A.A., Carrick F.N., Lavin M.F.;
RT "Remarkable sequence relatedness in the DNA encoding the major outer
RT membrane protein of Chlamydia psittaci (koala type I) and Chlamydia
RT pneumoniae.";
RL Gene 138:139-142(1994).
CC -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMERANE PROTEIN. OUTER MEMBRANE.
DR EMBL; X72023; CAA50906.1; -.
DR InterPro; IPR000604; Chlamydia_OMP.
DR Pfam; PF01308; Chlamydia_OMP; 1.
DR PRINTS; PR01334; CHLAMIDIAOMP.
DR ProDom; PD001717; Chlamydia_OMP; 1.
KW Outer membrane; Transmembrane; Porin; Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 389 MAJOR OUTER MEMBRANE PROTEIN.
SQ SEQUENCE 389 AA; 41579 MW; 5DC50E85A6F4E50F CRC64;

Query Match 85.7%; Score 78; DB 2; Length 389;
Best Local Similarity 93.8%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TGSAAANYTTAVDRPN 17

```
          |||| ||||||||||||
Db      90 TGSATANYTTAVDRPN 105


RESULT   2
Q9X717
ID   Q9X717      PRELIMINARY;      PRT;   341 AA.
AC   Q9X717;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydophila abortus.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LW508;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
DR   EMBL; M73040; AAD29103.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   NON_TER     341    341
SQ   SEQUENCE   341 AA;  36762 MW;  B5933C9BF6AAF171 CRC64;


  Query Match             70.3%;  Score 64;  DB 2;  Length 341;
  Best Local Similarity   75.0%;  Pred. No. 0.0073;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           ||:|||||  |  ||||
Db      54 TGTAAANYKTPTDRPN 69


RESULT   3
O69306
ID   O69306      PRELIMINARY;      PRT;   352 AA.
AC   O69306;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 PROTEIN (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM234;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ004874; CAA06183.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  37868 MW;  0AEA9B1E099EED41 CRC64;


  Query Match             70.3%;  Score 64;  DB 2;  Length 352;
  Best Local Similarity   75.0%;  Pred. No. 0.0075;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           ||:|||||  |  ||||
Db      90 TGTAAANYKTPTDRPN 105


RESULT   4
O69307
ID   O69307      PRELIMINARY;      PRT;   352 AA.
AC   O69307;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 PROTEIN (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM326;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ004875; CAA06184.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  37854 MW;  33589C6D1137CCDB CRC64;


  Query Match             70.3%;  Score 64;  DB 2;  Length 352;
  Best Local Similarity   75.0%;  Pred. No. 0.0075;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           ||:|||||  |  ||||
Db      90 TGTAAANYKTPTDRPN 105


RESULT   5
O70050
ID   O70050      PRELIMINARY;      PRT;   352 AA.
AC   O70050;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PMD623, PM112, PM364;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ005615; CAA06622.1; -.
DR   EMBL; AJ005613; CAA06620.1; -.
DR   EMBL; AJ005614; CAA06621.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  37826 MW;  2F9D092492E462D4 CRC64;


  Query Match             70.3%;  Score 64;  DB 2;  Length 352;
  Best Local Similarity   75.0%;  Pred. No. 0.0075;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGSAAANYTTAVDRPN 17
           ||:|||||  |  ||||
```

```
Db     90 TGTAAANYKTPTDRPN 105


RESULT   6
O70085
ID   O70085      PRELIMINARY;      PRT;   352 AA.
AC   O70085;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PMSH1, PM225;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ005618; CAA06625.1; -.
DR   EMBL; AJ005617; CAA06624.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  37854 MW;  391914AD146072CB CRC64;


  Query Match             70.3%;  Score 64;  DB 2;  Length 352;
  Best Local Similarity   75.0%;  Pred. No. 0.0075;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||:||||| |  ||||
Db     90 TGTAAANYKTPTDRPN 105


RESULT   7
O69305
ID   O69305      PRELIMINARY;      PRT;   353 AA.
AC   O69305;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OMP1 PROTEIN (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OCLH196;
RA   Hoelzle L.E., Steinhausen G., Eggemann G., Schiller I.,
RA   Wittenbrink M.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ004873; CAA06182.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER     353    353
SQ   SEQUENCE   353 AA;  37933 MW;  AC7D8FD9FA6E1728 CRC64;


  Query Match             70.3%;  Score 64;  DB 2;  Length 353;
  Best Local Similarity   75.0%;  Pred. No. 0.0076;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||:||||| |  ||||
Db     90 TGTAAANYKTPTDRPN 105


RESULT   8
Q9AIK1
ID   Q9AIK1      PRELIMINARY;      PRT;   388 AA.
AC   Q9AIK1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VS225;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269259; AAK00240.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     19       POTENTIAL.
FT   CHAIN        20    388       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   388 AA;  41573 MW;  8E232D22C9B9948D CRC64;


  Query Match             68.1%;  Score 62;  DB 2;  Length 388;
  Best Local Similarity   75.0%;  Pred. No. 0.018;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||||||:| |  ||||
Db     87 TGSAAADYKTPTDRPN 102


RESULT   9
Q9APM4
ID   Q9APM4      PRELIMINARY;      PRT;   389 AA.
AC   Q9APM4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN   OMP1.
OS   Chlamydophila abortus.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LLG;
RX   MEDLINE=20569239; PubMed=11119563;
RA   Vretou E., Psarrou E., Kaisar M., Vlisidou I., Salti-Montesanto V.,
RA   Longbottom D.;
RT   "Identification of protective epitopes by sequencing of the major
RT   outer membrane protein gene of a variant strain of Chlamydia psittaci
RT   serotype 1.";
RL   Infect. Immun. 69:607-612(2001).
DR   EMBL; AF272945; AAG53881.1; -.
KW   Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41897 MW;  20513C69C7DBAAF5 CRC64;


  Query Match             64.8%;  Score 59;  DB 2;  Length 389;
  Best Local Similarity   68.8%;  Pred. No. 0.059;
  Matches   11;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||:|||:| |  ||||
Db     90 TGTAAADYKTPTDRPN 105
```

```
RESULT  10
Q99V50
ID   Q99V50      PRELIMINARY;      PRT;   557 AA.
AC   Q99V50;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MENAQUINONE BIOSYNTHESIS PROTEIN.
GN   MEND OR SA0896.
OS   Staphylococcus aureus subsp. aureus N315.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=158879;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA   Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA   Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA   Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA   Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA   Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
RT   aureus.";
RL   Lancet 357:1225-1240(2001).
DR   EMBL; AP003132; BAB42141.1; -.
KW   Complete proteome.
SQ   SEQUENCE   557 AA;  63091 MW;  B38C5DA274972483 CRC64;


  Query Match           52.7%;  Score 48;  DB 2;  Length 557;
  Best Local Similarity  39.4%;  Pred. No. 6.2;
  Matches   13;  Conservative    3;  Mismatches    1;  Indels   16;  Gaps    2;

Qy      1 CT-GSAAANYTTAV--------------DRPN 17
          || |:|||||| |:                |||:
Db     78 CTSGTAAANYTPAIAESQISRIPLIVLTSDRPH 110


RESULT  11
Q9UFR7
ID   Q9UFR7      PRELIMINARY;      PRT;   662 AA.
AC   Q9UFR7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL 71.0 KDA PROTEIN.
GN   DKFZP434O051.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Poustka A., Klein M., Mewes H.W., Gassenhuber J., Wiemann S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL117400; CAB55901.1; -.
DR   InterPro; IPR000719; Euk_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Hypothetical protein; Transferase.
SQ   SEQUENCE   662 AA;  71003 MW;  8756E82919F6093D CRC64;


  Query Match           52.7%;  Score 48;  DB 4;  Length 662;
  Best Local Similarity  56.2%;  Pred. No. 7.4;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CTGSAAANYTTAVDRP 16
            ||||::| |  | | |
Db    251 CTGSSSACYALATDLP 266


RESULT  12
Q14148
ID   Q14148      PRELIMINARY;      PRT;  1241 AA.
AC   Q14148;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   KIAA0135 PROTEIN (FRAGMENT).
GN   KIAA0135.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127530; PubMed=8590280;
RA   Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. IV.
RT   The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 2:167-174(1995).
RN   [2]
RP   SEQUENCE OF 1-54 FROM N.A.
RA   Ceulemans H., Van Eynde A., Perez-Callejon E., Stalmans W., Bollen M.;
RT   "Structure and splice products of the human gene encoding sds22, a
RT   putative mitotic regulator of protein phosphatase-1.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   EMBL; D50925; BAA09484.1; -.
DR   EMBL; AF067137; AAC23506.1; -.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR000014; PAS.
DR   InterPro; IPR002290; Ser_thr_kin_actsite.
DR   InterPro; IPR001245; Tyr_kin.
DR   Pfam; PF00989; PAS; 3.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   SMART; SM00091; PAS; 2.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;
KW   Transferase.
FT   NON_TER       1       1
SQ   SEQUENCE   1241 AA;  134103 MW;  B651937986664A84 CRC64;


  Query Match           52.7%;  Score 48;  DB 4;  Length 1241;
  Best Local Similarity  56.2%;  Pred. No. 14;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CTGSAAANYTTAVDRP 16
            ||||::| |  | | |
Db    619 CTGSSSACYALATDLP 634


RESULT  13
Q9K5C5
ID   Q9K5C5      PRELIMINARY;      PRT;   326 AA.
AC   Q9K5C5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER MEMBRANE PROTEIN 1 (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=R54;
RX   PubMed=10919838;
RA   Herrmann B., Rahman R., Bergstroem S., Bonnedahl J., Olsen B.;
RT   "Chlamydophila abortus in a Brown Skua (Catharacta antarctica
RT   lonnbergi) from a Subantarctic Island.";
RL   Appl. Environ. Microbiol. 66:3654-3656(2000).
DR   EMBL; AJ243525; CAB96859.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   PRINTS; PR01334; CHLAMIDIAOMP.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   NON_TER     326    326
SQ   SEQUENCE   326 AA;  35345 MW;  6C5A20C8913743C8 CRC64;


  Query Match             51.6%;  Score 47;  DB 2;  Length 326;
  Best Local Similarity   62.5%;  Pred. No. 5.3;
  Matches   10;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          |||||:|  | || |
Db     36 TGSAAQDYKAAEDRAN 51


RESULT  14
Q9AIH9
ID   Q9AIH9      PRELIMINARY;      PRT;   389 AA.
AC   Q9AIH9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN   OMPA.
OS   Chlamydophila caviae.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83557;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GUINEA PIG INCLUSION CONJUNCTIVITIS, GPIC, ATCC VR813;
RX   MEDLINE=89212917; PubMed=2707861;
RA   Zhang Y.X., Morrison S.G., Caldwell H.D., Baehr W.;
RT   "Cloning and sequence analysis of the major outer membrane protein
RT   genes of two Chlamydia psittaci strains.";
RL   Infect. Immun. 57:1621-1625(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GUINEA PIG INCLUSION CONJUNCTIVITIS, GPIC, ATCC VR813;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269282; AAK00263.1; -.
KW   Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41932 MW;  2527A820C76F8310 CRC64;


  Query Match             51.6%;  Score 47;  DB 2;  Length 389;
  Best Local Similarity   56.2%;  Pred. No. 6.3;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TGSAAANYTTAVDRPN 17
          ||:|||:: |  || |
Db     89 TGNAAADFKTVADRNN 104


RESULT  15
Q9K5N9
ID   Q9K5N9      PRELIMINARY;      PRT;   168 AA.
AC   Q9K5N9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SINGLE-STRAND DNA-BINDING PROTEIN (PHAGE-RELATED PROTEIN).
GN   SSB OR BH4049.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001520; BAB07768.1; -.
DR   InterPro; IPR000424; SSB.
DR   Pfam; PF00436; SSB; 1.
DR   PROSITE; PS00735; SSB_1; 1.
KW   DNA-binding; Complete proteome.
SQ   SEQUENCE   168 AA;  18220 MW;  0D7C702E656232F6 CRC64;


  Query Match             50.5%;  Score 46;  DB 2;  Length 168;
  Best Local Similarity   64.3%;  Pred. No. 3.9;
  Matches    9;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 GSAAANYTTAVDRP 16
          | | ||:| ||:||
Db     23 GVAVANFTLAVNRP 36


Search completed: March 26, 2002, 13:40:11
Job time: 225 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:38:45 ; Search time 81.51 Seconds
(without alignments)
15.449 Million cell updates/sec

Title: US-09-709-201-96
Perfect score: 85
Sequence: 1 CASGTASNTTVAADRSN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	85	100.0	17	20	AAW95323	Costant and variab
2	46	54.1	163	21	AAY94712	Tumour necrosis fa
3	46	54.1	165	21	AAB00014	Peptide fragment o
4	46	54.1	183	16	AAR77421	BamTP delta53 nerv
5	46	54.1	198	21	AAY94720	Human type 2 tumou
6	46	54.1	225	21	AAY77463	Primate protein se
7	46	54.1	227	22	AAB66981	Tnfr2 protein. Un
8	46	54.1	235	19	AAW59665	Human soluble tumo
9	46	54.1	235	19	AAW52270	Tumour necrosis fa
10	46	54.1	235	20	AAW89234	Tumour necrosis in
11	46	54.1	235	21	AAY54440	Amino acid sequenc
12	46	54.1	235	21	AAY54441	Amino acid sequenc
13	46	54.1	235	21	AAY54442	A K108R/K120R muta
14	46	54.1	235	21	AAY54443	Wild type N-termna
15	46	54.1	235	22	AAB37685	Human 40 kDa TNF i
16	46	54.1	248	21	AAY94718	Human type 2 tumou
17	46	54.1	392	12	AAR11605	Human 75kD TNF-bin
18	46	54.1	392	20	AAY30935	Human tumour necro
19	46	54.1	461	12	AAR11001	40kD TNF inhibitor
20	46	54.1	461	12	AAR11141	Human TNF-R deduce
21	46	54.1	461	14	AAR42058	Fibroblast derived
22	46	54.1	461	16	AAR72504	p75 Tumour Necrosi
23	46	54.1	461	21	AAB37801	Human tumour necro
24	46	54.1	461	21	AAB18717	A human tumour nec
25	46	54.1	461	21	AAB01342	Death receptor. H
26	46	54.1	461	22	AAB35331	Human TNF receptor
27	46	54.1	461	22	AAB36698	Human tumour necro
28	46	54.1	461	22	AAB37686	Human 40 kDa TNF i
29	46	54.1	485	13	AAR24016	Fusion protein TNF
30	46	54.1	518	15	AAR51003	Sequence of a reco
31	46	54.1	518	22	AAB70001	sTNFR(075):Fc fusi
32	46	54.1	518	22	AAB50080	TNFR:Fc fusion pro
33	43	50.6	17	20	AAW95320	Costant and variab
34	43	50.6	287	22	AAB94116	Human protein sequ
35	43	50.6	427	21	AAB43107	Human ORFX ORF2871
36	42	49.4	389	21	AAY59209	B. halodurans clon
37	41	48.2	227	21	AAY77462	Rodent protein seq
38	41	48.2	258	22	AAB50082	Rat TNFR (p80) ext
39	41	48.2	474	12	AAR11142	TNF-R deduced from
40	41	48.2	487	22	AAB50084	TNFR:Fc fusion pro
41	40	47.1	213	22	AAB62541	B. melitensis viru
42	40	47.1	604	21	AAB27324	B. vulgaris NIM1 h
43	40	47.1	802	16	AAR70111	TBPII-GBP 130 fusi
44	40	47.1	802	16	AAR70112	TNF-R-GBP 130 fusi
45	40	47.1	826	21	AAG38474	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW95323
ID AAW95323 standard; Protein; 17 AA.
XX
AC AAW95323;
XX
DT 15-MAR-1999 (first entry)
XX
DE Costant and variable domain sequence of C. psitacci CPS92-106.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia psitacci.
XX
PN WO9850074-A2.
XX
PD 12-NOV-1998.
XX
PF 06-MAY-1998; 98WO-US09237.
XX
PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
PR 18-FEB-1998; 98US-0025176.
XX

PA (UYVA-) UNIV VANDERBILT.
XX
PI Mitchell WM, Stratton CW;
XX
DR WPI; 1999-059653/05.
XX
PT Composition with two agents effective against different stages of
PT chlamydial life cycle - comprises agent targetted against cryptic
PT phase, against elementary body phase, against replicating phase,
PT probenicid and antiporphyric
XX
PS Claim 4; Fig 3; 138pp; English.
XX
CC The invention relates to the diagnosis and management of infections by
CC Chlamydia species. The invention provides a composition that comprises
CC at least two agents, where each of the agents is effective against a
CC different phase of the chlamydial life cycle. The agents are selected
CC from: (a) agents targetted against cryptic phase of chlamydial life
CC cycle; (b) agents targetted against elementary body phase of chlamydial
CC life cycle; (c) agents targetted against replicating phase of chlamydial
CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC is used to elicit a protective immune response to Chlamydia infection in
CC an animal or human and is applied until the animal or human tests
CC negative for Chlamydia infection. It is also used to treat biological
CC material infected with Chlamydia. Diagnostic kits for antibody assays
CC against recombinant major outer membrane protein (MOMP), and for DNA
CC amplification assays for chlamydial genes, are used to diagnose disease,
CC e.g. autoimmune disease, an inflammatory disease or a disease that
CC occurs in an immuno-compromised individual, associated with Chlamydia
CC infection. The kits are used to detect chlamydial elementary bodies in a
CC sample. They are also used to monitor and/or modify the course of therapy
CC in a patient. The treatment reduces the acellular load of infectious
CC Ebstein Barr virus. The method is also used to treat porphyria, by
CC reducing the number of elementary bodies and applying a drug, e.g.
CC cimetidine, and antioxidants, to reduce the adverse effects associated
CC with porphyria. Sequences AAW95320 to AAW95323 represent constant and
CC variable domain sequences of various Chlamydia species.
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 85; DB 20; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAADRSN 17
||||||||||||||||| |
Db 1 casgtasnttvaadrsn 17

RESULT 2
AAY94712
ID AAY94712 standard; Protein; 163 AA.
XX
AC AAY94712;
XX
DT 29-JAN-2001 (first entry)
XX
DE Tumour necrosis factor receptor (TNFR) domain of TNFR-II.
XX
KW Tumour necrosis factor-receptor related protein; TR2; human; cancer;
KW chromosome p36.2-p36.3; arthritis; inflammation; autoimmune disease;
KW immunodeficiency; metastasis; haemolytic anaemia; asthma; X-linked SCID;
KW severely combined immunodeficiency; apoptosis inhibition;
KW Alzheimer's disease; Parkinson's disease; Crohn's disease.
XX
OS Homo sapiens.
XX
PN WO200056405-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US07521.
XX
PR 22-MAR-1999; 99US-0125683.
PR 26-MAR-1999; 99US-0126522.
PR 20-MAY-1999; 99US-0135169.
PR 06-AUG-1999; 99US-0147383.
XX
PA (NIJJ/) NI J.
PA (ROSE/) ROSEN C A.
PA (GENT/) GENTZ R L.
XX
PI Ni J, Rosen CA, Gentz RL;
XX
DR WPI; 2000-594519/56.
XX
PT Nucleic acid molecule encoding a human tumor necrosis factor receptor 2
PT and its two splice variants, useful for treating arthritis or
PT inflammation, cancer (such as follicular lymphomas) and
PT immunodeficiency disorders -
XX
PS Disclosure; Fig 16; 373pp; English.
XX
CC This invention relates to an isolated nucleic acid molecule encoding a
CC human tumour necrosis factor(TNF)-receptor related protein TR2. Included
CC in the invention are the two splice variants of TR2, TR2-SV1 and TR2-SV2.
CC The TR2 gene is located on chromosome 1 at position p36.2-p36.3. TR2 is a
CC member of the TNFR superfamily. The invention includes a method for the
CC treatment of arthritis or inflammation using an antibody directed against
CC a fragment of the TR2 protein. TR2 its agonists, antagonists and
CC antibodies exhibit cytostatic, dermatological, antianaemic,
CC immunosuppressive, antiallergic, antiarthritic, antiasthmatic,
CC antiinflammatory, neuroprotective, nootropic, antiparkinsonian, and
CC cerebroprotective activity. The methods are useful for treating arthritis
CC or inflammation, cancer (such as follicular lymphomas, carcinoma with p53
CC mutations, cardiac tumours, pancreatic, breast, or prostate cancer), an
CC immunodeficiency or for enhancing an in vivo leukocyte response to an
CC antigen. Anti-TR2 antibodies are useful for treating, inhibiting or
CC preventing autoimmune diseases (such as autoimmune haemolytic anaemia,
CC dermatitis, allergic encephalomyelitis, rheumatoid arthritis, asthma, and
CC inflammatory myopathies) and immunodeficiency disorders (such as severely
CC combined immunodeficiency (SCID)-X linked, B cell lymphoproliferative
CC disorder, or Nezelof syndrome-combined immunodeficiency with Igs). TR2,
CC TR2-SV1 and/or TR2-SV2 polynucleotides and polypeptides, agonists or
CC antagonists are useful for treating or preventing autoimmune diseases and
CC inhibit the growth, progression and/or metastasis of cancers. They are
CC also used to activate, differentiate or proliferate cancerous cells or
CC tissues, and can be used to treat diseases associated with increased cell
CC survival, or the inhibition of apoptosis, e.g. Alzheimer's disease,
CC Parkinson's disease, or Crohn's disease. The TR2 polypeptides are useful
CC as sources for generating antibodies, as molecular weight markers.
CC This sequence represents the tumour necrosis factor receptor (TNFR)
CC domain of the human TNFR-II protein. The sequence was used in the
CC characterisation of the TR2 receptor protein of the invention.
XX
SQ Sequence 163 AA;

Query Match 54.1%; Score 46; DB 21; Length 163;
Best Local Similarity 64.3%; Pred. No. 4.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAAD 14
|| || |||| : |
Db 126 capgtfsnttsstd 139

RESULT 3
AAB00014
ID AAB00014 standard; Peptide; 165 AA.
XX
AC AAB00014;
XX
DT 20-OCT-2000 (first entry)
XX

DE Peptide fragment of TNFR2.
XX
KW Tumour necrosis factor receptor homologue; TRH1; TNF; arthritis;
KW transplant rejection; activation; proliferation; differentiation;
KW apoptosis; immunosuppression; antiinflammatory; immunostimulation;
KW probe; primer; human.
XX
OS Homo sapiens.
XX
PN WO200034294-A2.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US29400.
XX
PR 11-DEC-1998; 98US-0111826.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Bowen MA, Siemers N;
XX
DR WPI; 2000-423364/36.
XX
PT Novel tumor necrosis factor receptor homologue-1 useful as a target for
PT immunosuppressive, antiinflammatory and/or immunostimulatory drug
PT development
XX
PS Disclosure; Fig 5; 42pp; English.
XX
CC The tumour necrosis factor receptor homologue TRH1 can be used for
CC treating a mammal e.g. a human, at risk for a disorder characterized
CC by an aberrant or unwanted level or biological activity of TRH1,
CC e.g. rheumatoid arthritis and transplant rejection. TRH1 may also be
CC useful to leach out or block a ligand which is found to bind to the
CC TRH1. TRH1 may be used in various drug screening techniques and to
CC identify fragments and analogs of a protein or peptide (agonist or
CC antagonist) which bind to TRH1. The TRH1 protein plays a role in
CC cellular function, cell activation, proliferation, differentiation,
CC and apoptosis. The interaction between the novel TNFr protein of the
CC present invention and intracellular signaling molecules and/or its
CC potential co-receptor may serve as a novel target for
CC immunosuppressive, antiinflammatory and/or immunostimulatory drug
CC development. Gene constructs can also be used as part of a gene
CC therapy protocol to deliver nucleic acids encoding the TRH1, or an
CC agonist or antagonist form of a TRH1 protein or peptide. Antibody
CC directed against TRH1 can be used to reject TRH1 in tissues
CC and cells. They can also be used to make targeted antibody that
CC destroy TRH1 expressing cells. Fragments of the TRH1 gene can be
CC used as diagnostic probes or as PCR primers. Fragments of the full
CC length gene may be used as hybridization probes for a cDNA library to
CC isolate the full length gene and to isolate other genes which have a
CC high sequence similarity. The probes may be used to identify a cDNA
CC clone corresponding to a full length transcript and a genomic clone
CC or clones that contain the complete gene including regulatory and
CC promoter regions, exons, and introns. This peptide fragment
CC corresponds to amino acids 38-202 of TNFR2 was used in a
CC homology comparison with TRH1. This cysteine rich motif was aligned
CC with the extracellular region of TRH1 (See GENESEQ record AAB00012).
XX
SQ Sequence 165 AA;

Query Match 54.1%; Score 46; DB 21; Length 165;
Best Local Similarity 64.3%; Pred. No. 4.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     127 capgtfsnttsstd 140
```

RESULT 4
AAR77421
ID AAR77421 standard; Protein; 183 AA.
XX
AC AAR77421;
XX
DT 10-FEB-1996 (first entry)
XX
DE BamTP delta53 nerve growth factor sequence.
XX
KW Nerve growth factor; neurotrophic factor; therapeutic;
KW protein refolding; NGF; plasmid pT3XI-2.
XX
OS Synthetic.
XX
PN WO9530686-A1.
XX
PD 16-NOV-1995.
XX
PF 02-MAY-1995; 95WO-US05423.
XX
PR 27-JUN-1994; 94US-0266080.
PR 09-MAY-1994; 94US-0240122.
XX
PA (SYNT) SYNTEX-SYNERGEN NEUROSCIENCE JOINT VENTU.
XX
PI Bonam D, Kohno T, Lile J, Rosendahl MS;
XX
DR WPI; 1995-404080/51.
DR N-PSDB; AAT05443.
XX
PT Process for bacterial expression of recombinant neurotrophic factor
PT - useful for promoting the survival and maintaining phenotypic
PT differentiation of nerve and glial cells.
XX
PS Example 1; Page 36-37; 57pp; English.
XX
CC The synthetic nerve growth factor (NGF) gene isolated from Bam TP
CC delta 53 plasmid pT3XI-2 is designed to optimize codons for
CC expression in Escherichia coli as well as create unique sites for
CC subsequent cloning steps. The recombinant protein is solubilized
CC and sulfonylated and allowed to refold in the presence of PEG and
CC urea. Biologically active NGF, used for promoting the survival of
CC and maintaining the phenotypic differentiation of nerve and glial
CC cells, is isolated and purified. This method breaks incorrectly
CC formed disulphide bonds and allows refolding of the factor into
CC the correct tertiary structure required for maximum yield of full
CC active protein.
XX
SQ Sequence 183 AA;

Query Match 54.1%; Score 46; DB 16; Length 183;
Best Local Similarity 64.3%; Pred. No. 5.2;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     143 capgtfsnttsstd 156
```

RESULT 5
AAY94720
ID AAY94720 standard; Protein; 198 AA.
XX
AC AAY94720;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human type 2 tumour necrosis factor receptor protein fragment.
XX
KW Tumour necrosis factor-receptor related protein; TR2; human; cancer;
KW chromosome p36.2-p36.3; arthritis; inflammation; autoimmune disease;
KW immunodeficiency; metastasis; haemolytic anaemia; asthma; X-linked SCID;
KW severely combined immunodeficiency; apoptosis inhibition;

```
KW  Alzheimer's disease; Parkinson's disease; Crohn's disease.
XX
OS  Homo sapiens.
XX
PN  WO200056405-A2.
XX
PD  28-SEP-2000.
XX
PF  22-MAR-2000; 2000WO-US07521.
XX
PR  22-MAR-1999;   99US-0125683.
PR  26-MAR-1999;   99US-0126522.
PR  20-MAY-1999;   99US-0135169.
PR  06-AUG-1999;   99US-0147383.
XX
PA  (NIJJ/) NI J.
PA  (ROSE/) ROSEN C A.
PA  (GENT/) GENTZ R L.
XX
PI  Ni J,  Rosen CA,  Gentz RL;
XX
DR  WPI; 2000-594519/56.
XX
PT  Nucleic acid molecule encoding a human tumor necrosis factor receptor 2
PT  and its two splice variants, useful for treating arthritis or
PT  inflammation, cancer (such as follicular lymphomas) and
PT  immunodeficiency disorders -
XX
PS  Disclosure; Fig 8; 373pp; English.
XX
CC  This invention relates to an isolated nucleic acid molecule encoding a
CC  human tumour necrosis factor(TNF)-receptor related protein TR2. Included
CC  in the invention are the two splice variants of TR2, TR2-SV1 and TR2-SV2.
CC  The TR2 gene is located on chromosome 1 at position p36.2-p36.3. TR2 is a
CC  member of the TNFR superfamily. The invention includes a method for the
CC  treatment of arthritis or inflammation using an antibody directed against
CC  a fragment of the TR2 protein. TR2 its agonists, antagonists and
CC  antibodies exhibit cytostatic, dermatological, antianaemic,
CC  immunosuppressive, antiallergic, antiarthritic, antiasthmatic,
CC  antiinflammatory, neuroprotective, nootropic, antiparkinsonian, and
CC  cerebroprotective activity. The methods are useful for treating arthritis
CC  or inflammation, cancer (such as follicular lymphomas, carcinoma with p53
CC  mutations, cardiac tumours, pancreatic, breast, or prostate cancer), an
CC  immunodeficiency or for enhancing an in vivo leukocyte response to an
CC  antigen. Anti-TR2 antibodies are useful for treating, inhibiting or
CC  preventing autoimmune diseases (such as autoimmune haemolytic anaemia,
CC  dermatitis, allergic encephalomyelitis, rheumatoid arthritis, asthma, and
CC  inflammatory myopathies) and immunodeficiency disorders (such as severely
CC  combined immunodeficiency (SCID)-X linked, B cell lymphoproliferative
CC  disorder, or Nezelof syndrome-combined immunodeficiency with Igs). TR2,
CC  TR2-SV1 and/or TR2-SV2 polynucleotides and polypeptides, agonists or
CC  antagonists are useful for treating or preventing autoimmune diseases and
CC  inhibit the growth, progression and/or metastasis of cancers. They are
CC  also used to activate, differentiate or proliferate cancerous cells or
CC  tissues, and can be used to treat diseases associated with increased cell
CC  survival, or the inhibition of apoptosis, e.g. Alzheimer's disease,
CC  Parkinson's disease, or Crohn's disease. The TR2 polypeptides are useful
CC  as sources for generating antibodies, as molecular weight markers.
CC  This sequence represents a fragment of the type 2 human tumour necrosis
CC  factor receptor protein. The sequence is used in the characterisation of
CC  the TR2 receptor protein of the invention.
XX
SQ  Sequence   198 AA;

  Query Match             54.1%;  Score 46;  DB 21;  Length 198;
  Best Local Similarity   64.3%;  Pred. No. 5.7;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     164 capgtfsnttsstd 177
```

```
RESULT  6
AAY77463
ID  AAY77463 standard; Protein; 225 AA.
XX
AC  AAY77463;
XX
DT  05-JUN-2000  (first entry)
XX
DE  Primate protein sequence, SEQ ID NO:14.
XX
KW  Imune disorder; inflammation; allergy; immunosuppressant;
KW  antiarthritic; antirheumatoid; antiinflammatory; dermatological;
KW  antithyroid.
XX
OS  Primates.
XX
PN  WO200001817-A2.
XX
PD  13-JAN-2000.
XX
PF  06-JUL-1999;   99WO-US12366.
XX
PR  06-JUL-1998;   98US-0110938.
PR  13-JUL-1998;   98US-0114466.
PR  23-JUL-1998;   98US-0093897.
PR  12-AUG-1998;   98US-0132968.
PR  18-AUG-1998;   98US-0136214.
PR  11-SEP-1998;   98US-0099999.
XX
PA  (SCHE ) SCHERING CORP.
XX
PI  Bates EEM,  Lebecque SJE,  Murphy EE,  Mattson JD,  Gorman DM;
PI  Hedrick JA,  Wang L,  Zlotnik A,  Murgolo NJ,  Greene JR,  Johnston JA;
PI  Bazan JF,  Mahony D,  Lees EM;
XX
DR  WPI; 2000-171015/15.
XX
PT  New isolated mammalian genes, used to develop products for treating
PT  e.g. immune, inflammatory or allergic abnormalities, cancers or
PT  degenerative conditions -
XX
PS  Disclosure; Page 170-171; 218pp; English.
XX
CC  The invention relates to a number of primate and/or rodent proteins, and
CC  the genes which encode them. The invention encompasses human dendritic
CC  cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC  factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
CC  human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub
CC  12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC  properties of ligands for proteins comprising a leucine-rich motif
CC  (LRR); human cyclin E2; cDNAs encoding these proteins; and antibodies
CC  against these proteins. The proteins can be used for modulating the
CC  physiology or development of a cell. They can be used to mediate uptake
CC  of substrates (e.g., prostaglandin-like molecules), to modulate or
CC  mediate cellular interactions (e.g., induce or prevent trafficking,
CC  proliferation, or differentiation of cells), or are intracellular
CC  proteins which are important in various cellular processes such as the
CC  deubiquitination of proteins or cell cycle regulation. The products can
CC  be used for treating medical conditions such as immune, inflammatory or
CC  allergic disorders, or abnormal cellular proliferation, for example,
CC  cancers or degenerative conditions. They can be used to modulate immune
CC  responses in disease states e.g., autoimmune disorders, including
CC  rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
CC  autoimmune thyroiditis, as well as acute and chronic inflammatory
CC  responses in which T cell activation, expansion, and/or immunological T
CC  cell memory play an important role. Sequences AAY77463-Y77464,
CC  AAY77474-Y77475 and AAY77484 represent primate proteins of undefined
CC  function, AAY77462 and AAY77481 are rodent proteins of undefined
CC  function, and AAY77482 is an avian protein of undefined function. These
CC  sequences are given in the sequence listing but are not referred to in
CC  the remainder of the specification.
XX
```

SQ Sequence 225 AA;

Query Match 54.1%; Score 46; DB 21; Length 225;
Best Local Similarity 64.3%; Pred. No. 6.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAAD 14
|| || |||| : |
Db 164 capgtfsnttsstd 177

RESULT 7
AAB66981
ID AAB66981 standard; Protein; 227 AA.
XX
AC AAB66981;
XX
DT 19-APR-2001 (first entry)
XX
DE Tnfr2 protein.
XX
KW Bone loss; osteoprotegerin; OPG; rheumatoid arthritis; hyperalgesia;
KW multiple sclerosis; osteoporosis; osteomyelitis; asthma; inflammation;
KW systemic lupus erythematosus; graft-versus-host disease; septic shock;
KW acute pancreatitis; Alzheimer's disease; anorexia; atherosclerosis; pain;
KW coronary condition; myocardial infarction; cancer; diabetes; psoriasis;
KW endometriosis; fever; glomerulonephritis; inflammatory bowel disease;
KW ischaemia; Parkinson's disease.
XX
OS Unidentified.
XX
PN WO200103719-A2.
XX
PD 18-JAN-2001.
XX
PF 07-JUL-2000; 2000WO-US18667.
XX
PR 09-JUL-1999; 99US-0350670.
PR 09-DEC-1999; 99US-0457647.
XX
PA (AMGE-) AMGEN INC.
XX
PI Boyle WJ, Lacey DL, Calzone FJ, Chang M, Senaldi G;
XX
DR WPI; 2001-103031/11.
XX
PT Treating conditions leading to bone loss such as rheumatoid arthritis,
PT multiple sclerosis and asthma, comprises administering an
PT osteoprotegerin protein in conjunction with e.g. inhibitors of
PT interleukin and tumor necrosis factor alpha -
XX
PS Disclosure; Fig 2; 316pp; English.
XX
CC The present invention relates to a method for treating conditions leading
CC to bone loss. The method comprises administering a purified and isolated
CC osteoprotegerin (OPG) protein (AAF57836-AAF57838 and AAB66974-AAB66976)
CC in conjunction with other substances such as tumour necrosis factor-alpha
CC (TNF-alpha) inhibitors, interleukin (IL)-6, -8 and -18 inhibitors, ICE
CC modulators, fibroblast growth factor (FGF)1-10 modulators and/or platelet
CC activating factor (PAF) antagonists. The method is useful for treating
CC conditions leading to bone loss such as rheumatoid arthritis, multiple
CC sclerosis, osteoporosis, osteomyelitis and asthma. The method is also
CC useful for treating inflammation, systemic lupus erythematosus (SLE) and
CC graft-versus-host disease (GvHD). Other diseases that can be treated
CC include acute pancreatitis, Alzheimer's disease, anorexia,
CC atherosclerosis, coronary conditions (e.g. myocardial infarction),
CC cancer, diabetes, endometriosis, fever, glomerulonephritis, hyperalgesia,
CC inflammatory bowel disease, ischaemia, pain, Parkinson's disease,
CC psoriasis and septic shock. The present sequence was used in a sequence
CC homology comparison.
XX
SQ Sequence 227 AA;

Query Match 54.1%; Score 46; DB 22; Length 227;
Best Local Similarity 64.3%; Pred. No. 6.7;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAAD 14
|| || |||| : |
Db 164 capgtfsnttsstd 177

RESULT 8
AAW59665
ID AAW59665 standard; Protein; 235 AA.
XX
AC AAW59665;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human soluble tumour necrosis factor receptor type II.
XX
KW Human; tumour necrosis factor; TNF; TNF receptor type II;
KW inflammatory disease; leukaemia; TNF binding protein;
KW anti-inflammatory drug; methotrexatess.
XX
OS Homo sapiens.
XX
PN WO9824463-A2.
XX
PD 11-JUN-1998.
XX
PF 08-DEC-1997; 97WO-US22733.
XX
PR 09-JUL-1997; 97US-0052023.
PR 06-DEC-1996; 96US-0032587.
PR 23-JAN-1997; 97US-0036355.
PR 07-FEB-1997; 97US-0039315.
XX
PA (AMGE-) AMGEN INC.
XX
PI Bendele AM, Edwards CK, Sennello RM;
XX
DR WPI; 1998-333039/29.
DR N-PSDB; AAV41549.
XX
PT Treatment of acute or chronic inflammatory disease, e.g. leukaemia -
PT by administering tumour necrosis factor binding protein and at least
PT one additional anti-inflammatory drug, e.g. methotrexate
XX
PS Disclosure; Fig 2; 104pp; English.
XX
CC This is the amino acid sequence of the human tumour necrosis factor
CC receptor type II, used in the method of the invention involving the
CC treatment of acute or chronic inflammatory disease such as leukaemia
CC by administering tumour necrosis factor binding protein and at least
CC one additional anti-inflammatory drug, e.g. methotrexate.
XX
SQ Sequence 235 AA;

Query Match 54.1%; Score 46; DB 19; Length 235;
Best Local Similarity 64.3%; Pred. No. 6.9;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAAD 14
|| || |||| : |
Db 142 capgtfsnttsstd 155

RESULT 9
AAW52270
ID AAW52270 standard; Protein; 235 AA.
XX

```
AC   AAW52270;
XX
DT   29-JUN-1998  (first entry)
XX
DE   Tumour necrosis factor inhibitor.
XX
KW   Soluble tumour necrosis factor receptor; sTNFR; TNF-mediated disease;
KW   tumour necrosis factor binding protein; autoimmune disease; arthritis;
KW   adult respiratory distress syndrome; cachexia/anorexia; cancer; therapy;
KW   tumour necrosis factor inhibitor; Alzheimer's disease; TNBP.
XX
OS   Homo sapiens.
XX
PN   WO9801555-A2.
XX
PD   15-JAN-1998.
XX
PF   09-JUL-1997;   97WO-US12244.
XX
PR   04-MAR-1997;   97US-0039792.
PR   09-JUL-1996;   96US-0021443.
PR   06-DEC-1996;   96US-0032534.
PR   23-JAN-1997;   97US-0037737.
PR   07-FEB-1997;   97US-0039314.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Edwards CK,  Fisher EF,  Kieft GL;
XX
DR   WPI; 1998-101052/09.
DR   N-PSDB; AAV19802.
XX
PT   Truncated and soluble forms of tumour necrosis factor receptor -
PT   useful for treating diseases involving factor, e.g. arthritis and
PT   adult respiratory distress syndrome
XX
PS   Claim 3; Fig 8; 205pp; English.
XX
CC   This sequence is the human tumour necrosis factor inhibitor. The protein
CC   was used to make the truncated soluble tumour necrosis factor receptor
CC   (sTNFR) proteins of the invention. The truncated sTNFR proteins and
CC   tumour necrosis factor binding proteins (TNBP) are used to treat any
CC   TNF-mediated disease, e.g. arthritis, adult respiratory distress
CC   syndrome, cachexia/anorexia, cancer, chronic fatigue syndrome, graft
CC   rejection, Alzheimer's disease and other autoimmune diseases. Cells
CC   transformed with a vector containing DNA encoding the protein may be used
CC   for production of recombinant sTNFR, which may also be used for measuring
CC   the amount of sTNFR in samples and to raise antibodies against sTNFR.
CC   TNBP may also be used in preparation of therapeutic compositions for
CC   treating the above diseases. The sTNFR proteins are well suited to large
CC   scale production (since they lack the deamidation site in region 111-126,
CC   so are more stable in vivo); contain fewer disulphide bonds and fewer
CC   epitopes, making them less antigenic than full-length proteins.
XX
SQ   Sequence   235 AA;

  Query Match             54.1%;  Score 46;  DB 19;  Length 235;
  Best Local Similarity   64.3%;  Pred. No. 6.9;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    142 capgtfsnttsstd 155


RESULT  10
AAW89234
ID   AAW89234 standard; Protein; 235 AA.
XX
AC   AAW89234;
XX
DT   04-MAR-1999  (first entry)
XX
DE   Tumour necrosis inhibitor 40 kDa protein.
XX
KW   Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;
KW   OPG; chimeric; fusion; dimerisation domain; autoimmune disease;
KW   inflammation; apoptosis.
XX
OS   Homo sapiens.
XX
PN   WO9849305-A1.
XX
PD   05-NOV-1998.
XX
PF   29-APR-1998;   98WO-US08631.
XX
PR   01-MAY-1997;   97US-0850188.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Boyle WJ,  Wooden S;
XX
DR   WPI; 1999-034661/03.
DR   N-PSDB; AAV81733.
XX
PT   New chimeric osteoprotegerin polypeptides - contain the
PT   osteoprotegerin dimerisation domain and a heterologous sequence,
PT   useful to treat TNF and TNFR-mediated disorders
XX
PS   Disclosure; Fig 3; 92pp; English.
XX
CC   The present invention describes a chimeric polypeptide (A1), comprising
CC   an osteoprotegerin (OPG) dimerisation domain fused to a heterologous
CC   amino acid sequence. Also described are: (1) a multimer polypeptide
CC   comprising covalently associated A1 monomers; (2) an isolated nucleic
CC   acid encoding A1; (3) an expression vector comprising the nucleic acid
CC   sequence; and (4) a host cell transformed or transfected with the
CC   expression vector so that the nucleic acid is expressible. The products
CC   from the present invention are useful to treat a variety of disorders
CC   including those related to receptor binding. Compositions comprising
CC   tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras
CC   are used to treat TNF and TNFR-mediated disorders such as inflammation,
CC   autoimmune diseases and disorders related to excessive apoptosis. The
CC   chimeras are also useful for detecting molecules which interact with
CC   fused heterologous sequences to identify potential new receptors and
CC   ligands. The present sequence represents the TNF inhibitor 40 kDa
CC   protein.
XX
SQ   Sequence   235 AA;

  Query Match             54.1%;  Score 46;  DB 20;  Length 235;
  Best Local Similarity   64.3%;  Pred. No. 6.9;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    142 capgtfsnttsstd 155


RESULT  11
AAY54440
ID   AAY54440 standard; Protein; 235 AA.
XX
AC   AAY54440;
XX
DT   25-APR-2000  (first entry)
XX
DE   Amino acid sequence of a K108R mutant of soluble p75 TNF receptor.
XX
KW   p75 tumour necrosis factor receptor; mutant; PEG conjugated protein;
KW   polyethylene glycol conjugation; PEG conjugation; protein activity.
XX
OS   Homo sapiens.
```

```
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 108
FT                   /note= "wild type Lys replaced with Arg"
XX
PN   WO9967291-A2.
XX
PD   29-DEC-1999.
XX
PF   18-JUN-1999;   99WO-US13953.
XX
PR   22-JUN-1998;   98US-0102530.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Pettit DK;
XX
DR   WPI; 2000-160577/14.
DR   N-PSDB; AAZ45759.
XX
PT   Novel methods for site-specific protein modification by mutagenesis by
PT   replacing polyethylene glycol reacting sites  -
XX
PS   Claim 17; Page 29; 36pp; English.
XX
CC   The present sequence represents a N-terminal fragment of a mutant of
CC   the soluble tumour necrosis factor (TNF) receptor, where the wild
CC   type Lys residue at position 108 is replaced with Arg. Lys108 (and
CC   Lys120) make contact between the p75 receptor and ligand. These
CC   residues are also potential polyethylene glycol (PEG) conjugation
CC   sites. The wild type p75 TNF receptor protein was mutated and conjugated
CC   to PEG, using the method of the invention. The specification descries
CC   a method for conjugating proteins with PEG to result in
CC   PEG-conjugated proteins having little or no reduction in a desired
CC   activity. Specifically, one or more amino acid residues that are
CC   critical for protein bioactivity and which are capable of reacting
CC   with PEG sites are deleted, prior to conjugation of the protein to PEG.
CC   The methods provide PEG conjugated proteins that are more homogeneous
CC   and present in higher yields. Conjugation does not take place at amino
CC   acid residues that are critical to the proteins bioactivity, thus
CC   maintaining the activity of the protein. The methods are used to
CC   produce PEG conjugated proteins.
XX
SQ   Sequence   235 AA;


  Query Match             54.1%;  Score 46;  DB 21;  Length 235;
  Best Local Similarity   64.3%;  Pred. No. 6.9;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     142 capgtfsnttsstd 155


RESULT  12
AAY54441
ID   AAY54441 standard; Protein; 235 AA.
XX
AC   AAY54441;
XX
DT   25-APR-2000  (first entry)
XX
DE   Amino acid sequence of a K120R mutant of soluble p75 TNF receptor.
XX
KW   p75 tumour necrosis factor receptor; mutant; PEG conjugated protein;
KW   polyethylene glycol conjugation; PEG conjugation; protein activity.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 120
FT                   /note= "wild type Lys replaced with Arg"
XX
PN   WO9967291-A2.
XX
PD   29-DEC-1999.
XX
PF   18-JUN-1999;   99WO-US13953.
XX
PR   22-JUN-1998;   98US-0102530.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Pettit DK;
XX
DR   WPI; 2000-160577/14.
DR   N-PSDB; AAZ45760.
XX
PT   Novel methods for site-specific protein modification by mutagenesis by
PT   replacing polyethylene glycol reacting sites  -
XX
PS   Claim 16; Page 31; 36pp; English.
XX
CC   The present sequence represents a N-terminal fragment of a mutant of
CC   the soluble tumour necrosis factor (TNF) receptor, where the wild
CC   type Lys residue at position 120 is replaced with Arg. Lys120 (and
CC   Lys108) make contact between the p75 receptor and ligand. These
CC   residues are also potential polyethylene glycol (PEG) conjugation
CC   sites. The wild type p75 TNF receptor protein was mutated and conjugated
CC   to PEG, using the method of the invention. The specification descries
CC   a method for conjugating proteins with PEG to result in
CC   PEG-conjugated proteins having little or no reduction in a desired
CC   activity. Specifically, one or more amino acid residues that are
CC   critical for protein bioactivity and which are capable of reacting
CC   with PEG sites are deleted, prior to conjugation of the protein to PEG.
CC   The methods provide PEG conjugated proteins that are more homogeneous
CC   and present in higher yields. Conjugation does not take place at amino
CC   acid residues that are critical to the proteins bioactivity, thus
CC   maintaining the activity of the protein. The methods are used to
CC   produce PEG conjugated proteins.
XX
SQ   Sequence   235 AA;


  Query Match             54.1%;  Score 46;  DB 21;  Length 235;
  Best Local Similarity   64.3%;  Pred. No. 6.9;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     142 capgtfsnttsstd 155


RESULT  13
AAY54442
ID   AAY54442 standard; Protein; 235 AA.
XX
AC   AAY54442;
XX
DT   25-APR-2000  (first entry)
XX
DE   A K108R/K120R mutant of soluble p75 TNF receptor.
XX
KW   p75 tumour necrosis factor receptor; mutant; PEG conjugated protein;
KW   polyethylene glycol conjugation; PEG conjugation; protein activity.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 108
FT                   /note= "wild type Lys replaced with Arg"
FT   Misc-difference 120
```

```
FT                   /note= "wild type Lys replaced with Arg"
XX
PN   WO9967291-A2.
XX
PD   29-DEC-1999.
XX
PF   18-JUN-1999;   99WO-US13953.
XX
PR   22-JUN-1998;   98US-0102530.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Pettit DK;
XX
DR   WPI; 2000-160577/14.
DR   N-PSDB; AAZ54761.
XX
PT   Novel methods for site-specific protein modification by mutagenesis by
PT   replacing polyethylene glycol reacting sites  -
XX
PS   Claim 16; Page 33-34; 36pp; English.
XX
CC   The present sequence represents a N-terminal fragment of a mutant of the
CC   soluble tumour necrosis factor (TNF) receptor, where the wild type Lys
CC   residues at positions 108 and 120 are replaced with Arg. Lys120 and
CC   Lys108 make contact between the p75 receptor and ligand. These
CC   residues are also potential polyethylene glycol (PEG) conjugation
CC   sites. The wild type p75 TNF receptor protein was mutated and conjugated
CC   to PEG, using the method of the invention. The specification descries
CC   a method for conjugating proteins with PEG to result in
CC   PEG-conjugated proteins having little or no reduction in a desired
CC   activity. Specifically, one or more amino acid residues that are
CC   critical for protein bioactivity and which are capable of reacting
CC   with PEG sites are deleted, prior to conjugation of the protein to PEG.
CC   The methods provide PEG conjugated proteins that are more homogeneous
CC   and present in higher yields. Conjugation does not take place at amino
CC   acid residues that are critical to the proteins bioactivity, thus
CC   maintaining the activity of the protein. The methods are used to
CC   produce PEG conjugated proteins.
XX
SQ   Sequence   235 AA;


  Query Match             54.1%;  Score 46;  DB 21;  Length 235;
  Best Local Similarity   64.3%;  Pred. No. 6.9;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     142 capgtfsnttsstd 155


RESULT  14
AAY54443
ID   AAY54443 standard; Protein; 235 AA.
XX
AC   AAY54443;
XX
DT   25-APR-2000  (first entry)
XX
DE   Wild type N-termnal fragment of the soluble p75 TNF receptor.
XX
KW   p75 tumour necrosis factor receptor; mutant; PEG conjugated protein;
KW   polyethylene glycol conjugation; PEG conjugation; protein activity.
XX
OS   Homo sapiens.
XX
PN   WO9967291-A2.
XX
PD   29-DEC-1999.
XX
PF   18-JUN-1999;   99WO-US13953.
XX
PR   22-JUN-1998;   98US-0102530.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Pettit DK;
XX
DR   WPI; 2000-160577/14.
DR   N-PSDB; AAZ45762.
XX
PT   Novel methods for site-specific protein modification by mutagenesis by
PT   replacing polyethylene glycol reacting sites  -
XX
PS   Claim 16; Page 35-36; 36pp; English.
XX
CC   The present sequence represents a N-terminal fragment of the soluble
CC   tumour necrosis factor (TNF) receptor. The wild type Lys residues at
CC   positions 108 and 120 are replaced with Arg (see AAY54441-42). Lys120
CC   and Lys108 make contact between the p75 receptor and ligand. These
CC   residues are also potential polyethylene glycol (PEG) conjugation
CC   sites. The wild type p75 TNF receptor protein was mutated and conjugated
CC   to PEG, using the method of the invention. The specification descries
CC   a method for conjugating proteins with PEG to result in
CC   PEG-conjugated proteins having little or no reduction in a desired
CC   activity. Specifically, one or more amino acid residues that are
CC   critical for protein bioactivity and which are capable of reacting
CC   with PEG sites are deleted, prior to conjugation of the protein to PEG.
CC   The methods provide PEG conjugated proteins that are more homogeneous
CC   and present in higher yields. Conjugation does not take place at amino
CC   acid residues that are critical to the proteins bioactivity, thus
CC   maintaining the activity of the protein. The methods are used to
CC   produce PEG conjugated proteins.
XX
SQ   Sequence   235 AA;


  Query Match             54.1%;  Score 46;  DB 21;  Length 235;
  Best Local Similarity   64.3%;  Pred. No. 6.9;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     142 capgtfsnttsstd 155


RESULT  15
AAB37685
ID   AAB37685 standard; Protein; 235 AA.
XX
AC   AAB37685;
XX
DT   02-MAR-2001  (first entry)
XX
DE   Human 40 kDa TNF inhibitor.
XX
KW   TNF inhibitor; antiinflammatory; Tumour Necrosis Factor; interleukin;
KW   IL-1; inflammatory disease; degenerative disease; human; lymphotoxin.
XX
OS   Homo sapiens.
XX
PN   US6143866-A.
XX
PD   07-NOV-2000.
XX
PF   19-JAN-1995;   95US-0375242.
XX
PR   19-JUL-1990;   90US-0555274.
PR   09-JUL-1993;   93US-0090366.
PR   18-JUL-1989;   89US-0381080.
PR   11-DEC-1989;   89US-0450329.
PR   07-FEB-1990;   90US-0479661.
XX
PA   (AMGE-) AMGEN INC.
XX
```

PI Squires C, King MW, Hale KK, Brewer MT, Thompson RC;
PI Vanderslice RW, Vannice J, Kohno T;
XX
DR WPI; 2001-006443/01.
XX
PT Novel 30 kDa tumor necrosis factor inhibitor analog comprising a
PT non-native cysteine residue cross-linked with polyethylene glycol,
PT useful for treating inflammatory and degenerative diseases mediated by
PT TNF -
XX
PS Example 12; Fig 38; 82pp; English.
XX
CC The present invention relates to Tumour Necrosis Factor (TNF) inhibitors
CC (see AAB37676 and AAB37685), which have TNF inhibitory activity. The
CC novel TNF inhibitors of the present invention are useful as therapeutic
CC agents for inhibiting the activity of TNF and interleukin (IL-1), and
CC for treating inflammatory and degenerative diseases mediated by TNF. The
CC present sequence is 40 kDa TNF inhibitor. The 40 kDa TNF inhibitor can
CC inhibit both TNF alpha and beta (lymphotoxin).
XX
SQ Sequence 235 AA;

Query Match 54.1%; Score 46; DB 22; Length 235;
Best Local Similarity 64.3%; Pred. No. 6.9;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     142 capgtfsnttsstd 155
```

Search completed: March 26, 2002, 13:38:46
Job time: 140 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:41:27 ; Search time 37.72 Seconds
(without alignments)
10.142 Million cell updates/sec

Title: US-09-709-201-96
Perfect score: 85
Sequence: 1 CASGTASNTTVAADRSN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	54.1	39	1	US-08-050-319B-41	Sequence 41, Appl
2	46	54.1	39	2	US-08-465-982-41	Sequence 41, Appl
3	46	54.1	163	2	US-08-219-237B-5	Sequence 5, Appli
4	46	54.1	163	4	US-08-477-347-13	Sequence 13, Appl
5	46	54.1	163	4	US-08-476-862-4	Sequence 4, Appli
6	46	54.1	163	4	US-08-468-560C-5	Sequence 5, Appli
7	46	54.1	164	2	US-08-232-087A-9	Sequence 9, Appli
8	46	54.1	227	3	US-08-974-022-48	Sequence 48, Appl
9	46	54.1	227	4	US-08-795-445A-48	Sequence 48, Appl
10	46	54.1	227	4	US-08-795-447A-48	Sequence 48, Appl
11	46	54.1	227	4	US-08-974-186-48	Sequence 48, Appl
12	46	54.1	227	4	US-08-795-446B-48	Sequence 48, Appl
13	46	54.1	235	4	US-09-326-394-4	Sequence 4, Appli
14	46	54.1	461	1	US-08-385-229-2	Sequence 2, Appli
15	46	54.1	461	2	US-08-650-000-2	Sequence 2, Appli
16	46	54.1	461	4	US-09-042-785A-7	Sequence 7, Appli
17	46	54.1	461	4	US-08-477-347-3	Sequence 3, Appli
18	46	54.1	461	4	US-09-006-353A-4	Sequence 4, Appli
19	46	54.1	461	4	US-08-476-862-2	Sequence 2, Appli
20	46	54.1	461	6	5395760-2	Patent No. 5395760
21	46	54.1	486	1	US-08-243-010-1	Sequence 1, Appli
22	46	54.1	518	1	US-08-385-229-4	Sequence 4, Appli
23	41	48.2	474	2	US-08-650-000-4	Sequence 4, Appli
24	41	48.2	474	4	US-09-042-785A-8	Sequence 8, Appli
25	41	48.2	474	6	5395760-4	Patent No. 5395760
26	40	47.1	223	1	US-08-278-091-13	Sequence 13, Appl
27	40	47.1	223	1	US-08-483-859-13	Sequence 13, Appl
28	40	47.1	223	1	US-08-472-173-13	Sequence 13, Appl
29	40	47.1	223	2	US-08-487-167-13	Sequence 13, Appl
30	40	47.1	223	2	US-08-482-816-13	Sequence 13, Appl
31	40	47.1	223	2	US-08-296-149-13	Sequence 13, Appl
32	40	47.1	223	2	US-08-801-499-13	Sequence 13, Appl
33	40	47.1	223	2	US-08-615-271-13	Sequence 13, Appl
34	40	47.1	223	3	US-09-074-660-13	Sequence 13, Appl
35	40	47.1	223	3	US-09-074-659-13	Sequence 13, Appl
36	40	47.1	223	3	US-09-106-468-13	Sequence 13, Appl
37	40	47.1	223	4	US-09-106-466A-13	Sequence 13, Appl
38	40	47.1	223	4	US-09-106-467-13	Sequence 13, Appl
39	39.5	46.5	1038	4	US-09-541-782-4	Sequence 4, Appli
40	39	45.9	1319	2	US-08-290-731C-2	Sequence 2, Appli
41	39	45.9	1336	2	US-08-290-731C-6	Sequence 6, Appli
42	38	44.7	327	1	US-08-240-049B-13	Sequence 13, Appl
43	38	44.7	327	1	US-08-240-049B-14	Sequence 14, Appl
44	38	44.7	327	1	US-08-259-148A-15	Sequence 15, Appl
45	38	44.7	327	1	US-08-259-148A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-050-319B-41
; Sequence 41, Application US/08050319B
; Patent No. 5633145
; GENERAL INFORMATION:
; APPLICANT: M.Feldmann, P.W. Gray,
; APPLICANT: M.J.C. Turner, F.M Brennan
; TITLE OF INVENTION: Modified human TNFalpha (Tumor
; TITLE OF INVENTION: Necrosis Factor alpha) Receptor
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reed & Robbins
; STREET: 635 Bryant Street
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,319B
; FILING DATE: 10-May-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Robbins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5150-0030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-050-319B-41

Query Match 54.1%; Score 46; DB 1; Length 39;
Best Local Similarity 64.3%; Pred. No. 0.45;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db       2 CAPGTFSNTTSSTD 15
```

```
RESULT   2
US-08-465-982-41
; Sequence 41, Application US/08465982
; Patent No. 5863786
;  GENERAL INFORMATION:
;    APPLICANT:  M.Feldmann, P.W. Gray,
;    APPLICANT:  M.J.C. Turner, F.M Brennan
;    TITLE OF INVENTION:  Modified human TNFalpha (Tumor
;    TITLE OF INVENTION:  Necrosis Factor alpha) Receptor
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Reed & Robbins
;      STREET:  635 Bryant Street
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,982
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/050,319
;      FILING DATE:  10-May-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Robbins, Roberta L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  5150-0030
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-465-982-41


  Query Match             54.1%;  Score 46;  DB 2;  Length 39;
  Best Local Similarity   64.3%;  Pred. No. 0.45;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db       2 CAPGTFSNTTSSTD 15


RESULT   3
US-08-219-237B-5
; Sequence 5, Application US/08219237B
; Patent No. 5874546
;  GENERAL INFORMATION:
;    APPLICANT:  NAGATA, Shigekazu
;    APPLICANT:  ITOH, Naoto
;    APPLICANT:  YONEHARA, Shin
;    TITLE OF INVENTION:  DNA Coding for Human Cell Surface Antigen
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  James W. Hellwege
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/219,237B
;      FILING DATE:  28-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/872,129
;      FILING DATE:  22-APR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  James W. Hellwege
;      REGISTRATION NUMBER:  28,808
;      REFERENCE/DOCKET NUMBER:  516762
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-219-237B-5


  Query Match             54.1%;  Score 46;  DB 2;  Length 163;
  Best Local Similarity   64.3%;  Pred. No. 2.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     126 CAPGTFSNTTSSTD 139


RESULT   4
US-08-477-347-13
; Sequence 13, Application US/08477347
; Patent No. 6232446
;  GENERAL INFORMATION:
;    APPLICANT:  WALLACH, David
;    APPLICANT:  BIGDA, Jacek
;    APPLICANT:  BELETSKY, Igor
;    APPLICANT:  METT, Igor
;    TITLE OF INVENTION:  TNF LIGANDS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,347
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/115,685
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 106271
;      FILING DATE:  08-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, G. Kevin
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  WALLACH-10
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  163 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-477-347-13


  Query Match           54.1%;  Score 46;  DB 4;  Length 163;
  Best Local Similarity  64.3%;  Pred. No. 2.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    126 CAPGTFSNTTSSTD 139


RESULT   5
US-08-476-862-4
; Sequence 4, Application US/08476862
; Patent No. 6262239
;  GENERAL INFORMATION:
;    APPLICANT:  WALLACH, David
;    APPLICANT:  BIGDA, Jacek
;    APPLICANT:  BELETSKY, Igor
;    APPLICANT:  METT, Igor
;    APPLICANT:  ENGELMANN, Hartmut
;    TITLE OF INVENTION:  TNF INHIBITORS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,862
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 107267
;      FILING DATE:  12-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 94039
;      FILING DATE:  06-APR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 91229
;      FILING DATE:  06-AUG-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 90339
;      FILING DATE:  18-MAY-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  WALLACH=12A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  163 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-476-862-4


  Query Match           54.1%;  Score 46;  DB 4;  Length 163;
  Best Local Similarity  64.3%;  Pred. No. 2.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    126 CAPGTFSNTTSSTD 139


RESULT   6
US-08-468-560C-5
; Sequence 5, Application US/08468560C
; Patent No. 6270998
;  GENERAL INFORMATION:
;    APPLICANT:  NAGATA, Shigekazu
;    APPLICANT:  ITOH, Naoto
;    APPLICANT:  YONEHARA, Shin
;    TITLE OF INVENTION:  DNA CODING FOR HUMAN CELL SURFACE
;    TITLE OF INVENTION:  ANTIGEN
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH, LLP.
;      STREET:  P.O. BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,560C
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURPHY JR., GERLAD M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  20-4393P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-205-8000
;      TELEFAX:  703-205-8050
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  163 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-468-560C-5


  Query Match           54.1%;  Score 46;  DB 4;  Length 163;
  Best Local Similarity  64.3%;  Pred. No. 2.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    126 CAPGTFSNTTSSTD 139


RESULT   7
US-08-232-087A-9
; Sequence 9, Application US/08232087A
; Patent No. 5866372
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Stein, Harald
;    APPLICANT:  D rkop, Horst
;    APPLICANT:  Latza, Ute
;    TITLE OF INVENTION:  Lymphoid CD30-Antigen
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,087A
;      FILING DATE:  08-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  756-103P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  164 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..164
;      OTHER INFORMATION:  /note= "TNFR2, see Fig. 5"
US-08-232-087A-9


  Query Match             54.1%;  Score 46;  DB 2;  Length 164;
  Best Local Similarity   64.3%;  Pred. No. 2.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     127 CAPGTFSNTTSSTD 140


RESULT   8
US-08-974-022-48
; Sequence 48, Application US/08974022
; Patent No. 6015938
;  GENERAL INFORMATION:
;    APPLICANT:  Boyle, Willaim J.
;    APPLICANT:  Lacey, David L.
;    APPLICANT:  Calzone, Frank J.
;    APPLICANT:  Chang, Ming-Shi
;    TITLE OF INVENTION:  OSTEOPROTEGERIN
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,022
;      FILING DATE:  12-DEC-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/577,788
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
;      REFERENCE/DOCKET NUMBER:  A-378
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  227 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-974-022-48


  Query Match             54.1%;  Score 46;  DB 3;  Length 227;
  Best Local Similarity   64.3%;  Pred. No. 3.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     164 CAPGTFSNTTSSTD 177


RESULT   9
US-08-795-445A-48
; Sequence 48, Application US/08795445A
; Patent No. 6284485
;  GENERAL INFORMATION:
;    APPLICANT:  Boyle, Willaim J.
;    APPLICANT:  Lacey, David L.
;    APPLICANT:  Calzone, Frank J.
;    APPLICANT:  Chang, Ming-Shi
;    TITLE OF INVENTION:  OSTEOPROTEGERIN
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,445A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/577,788
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
;      REFERENCE/DOCKET NUMBER:  A-378
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  227 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  protein
US-08-795-445A-48


  Query Match             54.1%;  Score 46;  DB 4;  Length 227;
  Best Local Similarity   64.3%;  Pred. No. 3.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    164 CAPGTFSNTTSSTD 177


RESULT  10
US-08-795-447A-48
; Sequence 48, Application US/08795447A
; Patent No. 6284728
;  GENERAL INFORMATION:
;    APPLICANT:  Boyle, William J.
;    APPLICANT:  Lacey, David L.
;    APPLICANT:  Calzone, Frank J.
;    APPLICANT:  Chang, Ming-Shi
;    TITLE OF INVENTION:  Osteoprotegerin
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  One Amgen Center Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91362-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,447A
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
;      REFERENCE/DOCKET NUMBER:  A-378D2
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  227 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-795-447A-48


  Query Match             54.1%;  Score 46;  DB 4;  Length 227;
  Best Local Similarity   64.3%;  Pred. No. 3.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    164 CAPGTFSNTTSSTD 177


RESULT  11
US-08-974-186-48
; Sequence 48, Application US/08974186
; Patent No. 6284740
;  GENERAL INFORMATION:
;    APPLICANT:  Boyle, Willaim J.
;    APPLICANT:  Lacey, David L.
;    APPLICANT:  Calzone, Frank J.
;    APPLICANT:  Chang, Ming-Shi
;    TITLE OF INVENTION:  OSTEOPROTEGERIN
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,186
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/577,788
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
;      REFERENCE/DOCKET NUMBER:  A-378
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  227 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-974-186-48


  Query Match             54.1%;  Score 46;  DB 4;  Length 227;
  Best Local Similarity   64.3%;  Pred. No. 3.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    164 CAPGTFSNTTSSTD 177


RESULT  12
US-08-795-446B-48
; Sequence 48, Application US/08795446B
; Patent No. 6288032
;  GENERAL INFORMATION:
;    APPLICANT:  Boyle, Willaim J.
;    APPLICANT:  Lacey, David L.
;    APPLICANT:  Calzone, Frank J.
;    APPLICANT:  Chang, Ming-Shi
;    TITLE OF INVENTION:  OSTEOPROTEGERIN
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,446B
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/577,788
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
```

```
;      REFERENCE/DOCKET NUMBER:  A-378
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  227 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-795-446B-48


  Query Match             54.1%;  Score 46;  DB 4;  Length 227;
  Best Local Similarity   64.3%;  Pred. No. 3.3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     164 CAPGTFSNTTSSTD 177


RESULT  13
US-09-326-394-4
; Sequence 4, Application US/09326394
; Patent No. 6306820
;  GENERAL INFORMATION:
;    APPLICANT:  Bendele, Alison M.
;    APPLICANT:  Sennello, Regina M.
;    APPLICANT:  Edwards, Carl K.
;    TITLE OF INVENTION:  COMBINATION THERAPY USING A TNF BINDING
;    TITLE OF INVENTION:  PROTEIN FOR TREATING TNF-MEDIATED DISEASES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/326,394
;      FILING DATE:  08-DEC-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,587
;      FILING DATE:  06-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/036,355
;      FILING DATE:  23-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/039,315
;      FILING DATE:  07-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/052,023
;      FILING DATE:  09-JUL-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zindrick, Thomas K.
;      REGISTRATION NUMBER:  32,185
;      REFERENCE/DOCKET NUMBER:  A-430D
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  235 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-326-394-4


  Query Match             54.1%;  Score 46;  DB 4;  Length 235;
  Best Local Similarity   64.3%;  Pred. No. 3.4;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     142 CAPGTFSNTTSSTD 155


RESULT  14
US-08-385-229-2
; Sequence 2, Application US/08385229
; Patent No. 5605690
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Cindy A.
;    APPLICANT:  Smith, Craig A.
;    TITLE OF INVENTION:  Method of Treating TNF-Dependent
;    TITLE OF INVENTION:  Inflammation Using Tumor Necrosis Factor Antagonists
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/385,229
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/946,236
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wight, Christopher L.
;      REGISTRATION NUMBER:  31,680
;      REFERENCE/DOCKET NUMBER:  2503
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 587-0430
;      TELEFAX:  (206) 587-0606
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  461 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-385-229-2


  Query Match             54.1%;  Score 46;  DB 1;  Length 461;
  Best Local Similarity   64.3%;  Pred. No. 7.4;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     164 CAPGTFSNTTSSTD 177


RESULT  15
US-08-650-000-2
; Sequence 2, Application US/08650000
; Patent No. 5945397
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Craig A.
;    APPLICANT:  Goodwin, Raymond G.
;    APPLICANT:  Beckmann, M. Patricia
;    TITLE OF INVENTION:  Tumor Necrosis Factor Receptors
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
```

; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/650,000
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,453
; FILING DATE:
; APPLICATION NUMBER: US/08/038,765
; FILING DATE:
; APPLICATION NUMBER: US 403,241
; FILING DATE: 05-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 405,370
; FILING DATE: 11-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 421,417
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 523,635
; FILING DATE: 10-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Wight, Christopher L.
; REGISTRATION NUMBER: 31,680
; REFERENCE/DOCKET NUMBER: 2501-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-650-000-2

Query Match 54.1%; Score 46; DB 2; Length 461;
Best Local Similarity 64.3%; Pred. No. 7.4;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     164 CAPGTFSNTTSSTD 177
```

Search completed: March 26, 2002, 13:41:27
Job time: 301 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:37:18 ; Search time 42.75 Seconds
(without alignments)
30.292 Million cell updates/sec

Title: US-09-709-201-96
Perfect score: 85
Sequence: 1 CASGTASNTTVAADRSN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	76	89.4	392	2	A40371	major outer membra
2	46	54.1	461	1	A35356	tumor necrosis fac
3	46	54.1	1518	2	T28880	hypothetical prote
4	46	54.1	4753	1	A47437	LDL-receptor-relat
5	45	52.9	390	2	T25642	hypothetical prote
6	45	52.9	767	2	S55618	hypothetical prote
7	45	52.9	942	2	T13014	cytochrome b245 be
8	43	50.6	425	2	S41099	protein kinase (EC
9	43	50.6	427	2	T46265	hypothetical prote
10	43	50.6	1462	2	T06819	DNA topoisomerase
11	42	49.4	390	2	E84066	hypothetical prote
12	42	49.4	518	2	F70831	probable PPE prote
13	42	49.4	621	2	S35092	plakoglobin - mous
14	41	48.2	459	2	I48854	gene murine tumour
15	41	48.2	474	2	B38634	tumor necrosis fac
16	41	48.2	1237	2	H81660	DNA polymerase III
17	40	47.1	226	2	B41378	cytochrome c553i p
18	40	47.1	259	1	TRSMG	trypsin (EC 3.4.21
19	40	47.1	490	2	T49096	hypothetical prote
20	40	47.1	492	2	T47720	pyruvate kinase-li
21	40	47.1	510	2	T47704	pyruvate kinase-li
22	40	47.1	760	2	S18686	Sc/SvMr protein -
23	39	45.9	202	2	JX0228	multicatalytic end
24	39	45.9	205	2	S17522	multicatalytic end
25	39	45.9	226	2	T40487	proteasome compone
26	39	45.9	239	2	B54589	proteasome subunit
27	39	45.9	323	2	B83757	sodium-dependent t
28	39	45.9	527	2	T04329	importin alpha - t
29	39	45.9	644	2	T37800	probable lysophosp
30	39	45.9	710	2	E69665	nitrate reductase
31	39	45.9	1181	2	T30578	myosin IC - slime
32	39	45.9	1181	2	T19736	hypothetical prote
33	39	45.9	1336	2	S25716	Ras guanine nucleo
34	39	45.9	1975	2	B81192	hemagglutinin/hemo
35	39	45.9	1995	2	G81044	hemagglutinin/hemo
36	38	44.7	85	2	S60856	M protein precurso
37	38	44.7	97	2	S60846	M protein precurso
38	38	44.7	117	1	JQ1745	glycoprotein J - c
39	38	44.7	174	2	A86358	Similar to blue co
40	38	44.7	215	2	S61337	multicatalytic end
41	38	44.7	315	2	E81937	probable transmemb
42	38	44.7	315	2	B81168	transporter NMB070
43	38	44.7	334	2	T50139	GNS1/SUR4 family p
44	38	44.7	335	2	S71796	centrosome-binding
45	38	44.7	356	2	T27052	hypothetical prote

ALIGNMENTS

RESULT 1
A40371
major outer membrane protein precursor - Chlamydophila psittaci (strain Fpn/pring)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 31-Mar-2000
C;Accession: I40859; A40371; S16137
R;Storey, C.; Lusher, M.; Yates, P.; Richmond, S.
J. Gen. Microbiol. 139, 2621-2626, 1993
A;Title: Evidence for Chlamydia pneumoniae of non-human origin.
A;Reference number: I40739; MUID:94103736
A;Accession: I40859
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A;Molecule type: DNA
A;Residues: 1-392 <RES>
A;Cross-references: EMBL:X61096; NID:g40564; PIDN:CAA43409.1; PID:g40565
A;Experimental source: strain Fpn
C;Genetics:
A;Gene: MOMP
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-392/Product: major outer membrane protein #status predicted <MAT>

Query Match 89.4%; Score 76; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ASGTASNTTVAADRSN 17
||||||||||||||||
Db 92 ASGTASNTTVAADRSN 107

RESULT 2
A35356
tumor necrosis factor receptor 2 precursor [validated] - human
N;Alternate names: 75K tumor necrosis factor receptor; TNF receptor type 2
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Dec-2000
C;Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094
R;Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower,
Science 248, 1019-1023, 1990
A;Title: A receptor for tumor necrosis factor defines an unusual family of cellular a
A;Reference number: A35356; MUID:90260639
A;Accession: A35356
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <SMI>
A;Cross-references: GB:M32315; NID:g189185; PIDN:AAA59929.1; PID:g189186
R;Kohno, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squir
Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990
A;Title: A second tumor necrosis factor receptor gene product can shed a naturally oc

A;Reference number: A36475; MUID:91045991
A;Accession: A36475
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-195,'R',197-461 <KOH>
A;Cross-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758
R;Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, M.
Cytokine 2, 231-237, 1990
A;Title: Two human TNF receptors have similar extracellular, but distinct intracellular,
A;Reference number: A48416; MUID:91370690
A;Accession: A48416
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 23-461 <DEM>
A;Cross-references: GB:S63368; NID:g235648; PIDN:AAB19824.1; PID:g235649
A;Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBIP:63371)
R;Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990
A;Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demonstra
A;Reference number: A36007; MUID:90349572
A;Accession: A36007
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 116-140,'P',142-195,'R',197-362,'T',364-461 <HEL>
A;Cross-references: GB:M35857; NID:g339751; PIDN:AAA63262.1; PID:g339752
R;Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus, M.
J. Biol. Chem. 265, 20131-20138, 1990
A;Title: Purification and partial amino acid sequence analysis of two distinct tumor nec
A;Reference number: A23666; MUID:91056048
A;Accession: A23666
A;Status: preliminary
A;Molecule type: protein
A;Residues: 23-40;65-69;136-141;300-306 <LOE>
R;Engelmann, H.; Novick, D.; Wallach, D.
J. Biol. Chem. 265, 1531-1536, 1990
A;Title: Two tumor necrosis factor-binding proteins purified from human urine. Evidence
A;Reference number: A35010; MUID:90110215
A;Accession: B35010
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-31 <ENG>
R;Kuhnert, P.; Kemper, O.; Wallach, D.
Gene 150, 381-386, 1994
A;Title: Cloning, sequencing and partial functional characterization of the 5' region of
A;Reference number: I38094; MUID:95121934
A;Accession: I38094
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-37 <RES>
A;Cross-references: EMBL:X80021; NID:g666044; PIDN:CAA56324.1; PID:g825701
C;Genetics:
A;Gene: GDB:TNFR2
A;Cross-references: GDB:125914; OMIM:191191
A;Map position: 1p36.2-1p36.2
A;Introns: 26/3
A;Note: the list of introns is incomplete
C;Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C;Keywords: duplication; glycoprotein; receptor; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-416/Product: tumor necrosis factor receptor 2 #status experimental <MAT>
F;40-76/Domain: NGF receptor repeat homology <NG1>
F;78-119/Domain: NGF receptor repeat homology <NG2>
F;120-162/Domain: NGF receptor repeat homology <NG3>
F;164-201/Domain: NGF receptor repeat homology <NG4>
F;262-279/Domain: transmembrane #status predicted <TMN>
F;280-461/Domain: intracellular #status predicted <INT>
F;171,193/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 54.1%; Score 46; DB 1; Length 461;
Best Local Similarity 64.3%; Pred. No. 6.8;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAAD 14
|| || |||| : |
Db 164 CAPGTFSNTTSSTD 177

RESULT 3
T28880
hypothetical protein R08E3.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C;Accession: T28880
R;Hawkins, J.
submitted to the EMBL Data Library, December 1995
A;Description: The sequence of C. elegans cosmid R08E3.
A;Reference number: Z20536
A;Accession: T28880
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1518 <HAW>
A;Cross-references: EMBL:U41027; PIDN:AAC48194.1; GSPDB:GN00028; CESP:R08E3.1
A;Experimental source: strain Bristol N2; clone R08E3
C;Genetics:
A;Gene: CESP:R08E3.1
A;Map position: X
A;Introns: 41/3; 65/1; 90/1; 120/2; 166/3; 214/1; 244/3; 285/3; 341/1; 370/1; 400/3;
; 1167/2; 1198/2; 1215/2; 1270/1; 1300/3; 1332/1; 1350/1; 1394/2; 1429/2; 1461/3; 150
C;Superfamily: Caenorhabditis elegans hypothetical protein R08E3.1

Query Match 54.1%; Score 46; DB 2; Length 1518;
Best Local Similarity 52.9%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAADRSN 17
| | || :| ||||
Db 1275 CVYGLQSNIKIAIDRSN 1291

RESULT 4
A47437
LDL-receptor-related protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Aug-2000
C;Accession: A47437; S27801; T21547
R;Yochem, J.; Greenwald, I.
Proc. Natl. Acad. Sci. U.S.A. 90, 4572-4576, 1993
A;Title: A gene for a low density lipoprotein receptor-related protein in the nematod
A;Reference number: A47437; MUID:93281621
A;Accession: A47437
A;Molecule type: DNA
A;Residues: 1-4753 <YOC>
A;Cross-references: GB:M96150; NID:g156359; PIDN:AAA28105.1; PID:g156360
A;Note: nucleotide sequence not given; translation not complete in this paper
R;Yochem, J.; Greenwald, I.
submitted to the EMBL Data Library, July 1992
A;Description: A gene for an LDL receptor-related protein (LPR) in the nematode C.ele
A;Reference number: S27801
A;Accession: S27801
A;Molecule type: DNA
A;Residues: 1-4753 <YO2>
A;Cross-references: EMBL:M96150; NID:g156359; PIDN:AAA28105.1; PID:g156360
R;Wilkinson, J.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19439
A;Accession: T21547
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4753 <WIL>
A;Cross-references: EMBL:Z73907; PIDN:CAA98124.1; GSPDB:GN00019; CESP:F29D11.1
A;Experimental source: clone F29D11
C;Genetics:
A;Gene: LPR
A;Map position: 1

```
A;Introns: 31/1; 88/1; 132/1; 172/3; 219/1; 298/1; 463/2; 526/2; 585/3; 780/2; 874/2; 97
15/1
C;Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
C;Keywords: tandem repeat; transmembrane protein
F;53-87/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;92-131/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;138-175/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;182-218/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;223-257/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F;262-297/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F;302-336/Domain: EGF homology <EGF1>
F;1054-1095/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F;1101-1138/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F;1146-1182/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F;1187-1223/Domain: LDL receptor ligand-binding repeat homology <LD10>
F;1228-1263/Domain: LDL receptor ligand-binding repeat homology <LD11>
F;1270-1307/Domain: LDL receptor ligand-binding repeat homology <LD12>
F;1313-1350/Domain: LDL receptor ligand-binding repeat homology <LD13>
F;1359-1396/Domain: LDL receptor ligand-binding repeat homology <LD14>
F;1441-1475/Domain: EGF homology <EGF>
F;1611-1654/Domain: LDL receptor YWTD-containing repeat homology <YW33>
F;2792-2829/Domain: LDL receptor ligand-binding repeat homology <LD15>
F;2834-2868/Domain: LDL receptor ligand-binding repeat homology <LD16>
F;2874-2912/Domain: LDL receptor ligand-binding repeat homology <LD17>
F;2919-2956/Domain: LDL receptor ligand-binding repeat homology <LD18>
F;2961-2997/Domain: LDL receptor ligand-binding repeat homology <LD19>
F;3006-3044/Domain: LDL receptor ligand-binding repeat homology <LD20>
F;3049-3093/Domain: LDL receptor ligand-binding repeat homology <LD21>
F;3100-3135/Domain: LDL receptor ligand-binding repeat homology <LD22>
F;3140-3174/Domain: LDL receptor ligand-binding repeat homology <LD23>
F;3187-3222/Domain: LDL receptor ligand-binding repeat homology <LD24>
F;3586-3623/Domain: EGF homology <EGX1>
F;3627-3666/Domain: LDL receptor ligand-binding repeat homology <LD25>
F;3671-3705/Domain: LDL receptor ligand-binding repeat homology <LD26>
F;3709-3746/Domain: LDL receptor ligand-binding repeat homology <LD27>
F;3753-3788/Domain: LDL receptor ligand-binding repeat homology <LD28>
F;3793-3830/Domain: LDL receptor ligand-binding repeat homology <LD29>
F;3833-3871/Domain: LDL receptor ligand-binding repeat homology <LD30>
F;3878-3912/Domain: LDL receptor ligand-binding repeat homology <LD31>
F;3917-3951/Domain: LDL receptor ligand-binding repeat homology <LD32>
F;3959-3995/Domain: LDL receptor ligand-binding repeat homology <LD33>
F;4000-4040/Domain: LDL receptor ligand-binding repeat homology <LD34>
F;4049-4083/Domain: LDL receptor ligand-binding repeat homology <LD35>
F;4092-4130/Domain: EGF homology <EGF2>
F;4343-4386/Domain: LDL receptor YWTD-containing repeat homology <YW38>


  Query Match             54.1%;  Score 46;  DB 1;  Length 4753;
  Best Local Similarity   52.9%;  Pred. No. 64;
  Matches    9;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAADRSN 17
           ||:|   | ||| || :
Db    4056 CANGKCVNGTVACDRKD 4072


RESULT   5
T25642
hypothetical protein C46H11.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25642
R;Miller, N.; Bradshaw, H.; Wamsley, P.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid C46H11.
A;Reference number: Z20061
A;Accession: T25642
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-390 <MIL>
A;Cross-references: EMBL:U88314; PIDN:AAB42362.1; GSPDB:GN00019; CESP:C46H11.8
A;Experimental source: strain Bristol N2; clone C46H11
C;Genetics:
A;Gene: CESP:C46H11.8
A;Map position: 1
A;Introns: 26/1; 51/1; 189/1; 345/1


  Query Match             52.9%;  Score 45;  DB 2;  Length 390;
  Best Local Similarity   56.2%;  Pred. No. 8.5;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ASGTASNTTVAADRSN 17
             :| |:|::|  |||||
Db     182 SSTTSSSSTTCADRSN 197


RESULT   6
S55618
hypothetical protein 24 - equine herpesvirus 2
C;Species: equine herpesvirus 2
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999
C;Accession: S55618
R;Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A;Title: The DNA sequence of equine herpesvirus 2.
A;Reference number: S55594; MUID:95302501
A;Accession: S55618
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-767 <TEL>
A;Cross-references: GB:U20824; NID:g695172; PIDN:AAC13811.1; PID:g695196
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995


  Query Match             52.9%;  Score 45;  DB 2;  Length 767;
  Best Local Similarity   57.1%;  Pred. No. 16;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           |||||| |  :: |
Db     511 CASGTAINMNISGD 524


RESULT   7
T13014
cytochrome b245 beta chain homolog F8L21.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 29-Oct-1999
C;Accession: T13014
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.;
submitted to the Protein Sequence Database, July 1999
A;Reference number: Z17587
A;Accession: T13014
A;Molecule type: DNA
A;Residues: 1-942 <BEV>
A;Cross-references: EMBL:AL096882; GSPDB:GN00062; ATSP:F8L21.20
A;Experimental source: cultivar Columbia; BAC clone F8L21
C;Genetics:
A;Gene: ATSP:F8L21.20
A;Map position: 4
A;Introns: 211/1; 282/3; 322/3; 458/3; 587/3; 619/3; 658/2; 714/2; 767/3; 837/3; 880/


  Query Match             52.9%;  Score 45;  DB 2;  Length 942;
  Best Local Similarity   43.8%;  Pred. No. 20;
  Matches    7;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAADRS 16
           | ||: ||: :::| |
Db     769 CISGSCSNSNISSDHS 784


RESULT   8
S41099
protein kinase (EC 2.7.1.37), cAMP-dependent, catalytic chain C - fungus (Blastocladi
```

```
C;Species: Blastocladiella emersonii
C;Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 28-May-1999
C;Accession: S41099; S77889; S77890
R;Franco de Oliveira, J.C.; Cantisani Borges, A.C.; do Valle Marques, M.; Lopes Gomes, S
Eur. J. Biochem. 219, 555-562, 1994
A;Title: Cloning and characterization of the gene for the catalytic subunit of cAMP-depe
A;Reference number: S41099; MUID:94139736
A;Accession: S41099
A;Molecule type: DNA
A;Residues: 1-425 <FRA>
A;Cross-references: GB:L17008; NID:g304272; PIDN:AAA20074.1; PID:g304273
A;Accession: S77889
A;Molecule type: mRNA
A;Residues: 22-425 <FRB>
A;Cross-references: GB:M81709; GB:L17038; NID:g507140; PIDN:AAA19440.1; PID:g507141
A;Accession: S77890
A;Molecule type: protein
A;Residues: 2-16 <FRC>
C;Genetics:
A;Introns: 209/3; 243/3; 315/1
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cAMP binding; magnesium; phosphoprotein; phosphotransferase; serine/thr
F;2-425/Product: protein kinase, cAMP-dependent, catalytic chain C #status experimental
F;114-370/Domain: protein kinase homology <KIN>
F;122-130/Region: protein kinase ATP-binding motif
F;127,128,194,200,243,256/Binding site: Mg-ATP (Phe, Gly, Glu, Glu, Glu, Thr) #status pr
F;145,164,239,241/Active site: Lys, Glu, Asp, Lys #status predicted
F;244,257/Binding site: magnesium (Asn, Asp) #status predicted
F;270/Binding site: phosphate (Thr) (covalent) #status predicted


  Query Match             50.6%;  Score 43;  DB 2;  Length 425;
  Best Local Similarity   62.5%;  Pred. No. 19;
  Matches   10;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAADRSN 17
          || |||:|| ||   |
Db     40 ASSTASSTTTAAASGN 55


RESULT  9
T46265
hypothetical protein DKFZp761A052.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46265
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46265
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <AAA>
A;Cross-references: EMBL:AL137509
A;Experimental source: adult amygdala; clone DKFZp761A052
C;Genetics:
A;Note: DKFZp761A052.1


  Query Match             50.6%;  Score 43;  DB 2;  Length 427;
  Best Local Similarity   50.0%;  Pred. No. 20;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAADRS 16
          || ||:|  :  |||:
Db    347 CAPGTSSQFSAGADRA 362


RESULT  10
T06819
DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) II - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C;Accession: T06819
R;Reddy, M.K.; Nair, S.; Tewari, K.K.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z15832
A;Accession: T06819
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1462 <RED>
A;Cross-references: EMBL:Y14559; PIDN:CAA74891.1
C;Genetics:
A;Note: TOPII
C;Function:
A;Description: involved in DNA replication and chromosome condensation
C;Superfamily: eukaryotic type II DNA topoisomerase; phage T4 DNA topoisomerase (ATP-
C;Keywords: ATP; DNA binding; isomerase


  Query Match             50.6%;  Score 43;  DB 2;  Length 1462;
  Best Local Similarity   50.0%;  Pred. No. 64;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          |::|:|||| :: |
Db   1400 CSAGSASNTPLSED 1413


RESULT  11
E84066
hypothetical protein BH3333 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C;Accession: E84066
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20263314
A;Accession: E84066
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-390 <STO>
A;Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB07052.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3333


  Query Match             49.4%;  Score 42;  DB 2;  Length 390;
  Best Local Similarity   61.5%;  Pred. No. 26;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      5 TASNTTVAADRSN 17
          | || |: |||:|
Db    122 TVSNLTIDADRTN 134


RESULT  12
F70831
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: F70831
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A;Reference number: A70500; MUID:98295987
A;Accession: F70831
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-518 <COL>
A;Cross-references: GB:AL021932; GB:AL123456; NID:g3261527; PIDN:CAA17410.1; PID:e125
```

A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: PPE

Query Match 49.4%; Score 42; DB 2; Length 518;
Best Local Similarity 81.8%; Pred. No. 34;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SGTASNTTVAA 13
:||||| ||||
Db 474 AGTASNETVAA 484

RESULT 13
S35092
plakoglobin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-1999
C;Accession: S35092
R;Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.
Science 257, 1142-1144, 1992
A;Title: Plakoglobin and beta-catenin: distinct but closely related.
A;Reference number: S35091; MUID:92376536
A;Accession: S35092
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-621 <BUT>
A;Cross-references: EMBL:M90365
C;Keywords: cytoskeleton

Query Match 49.4%; Score 42; DB 2; Length 621;
Best Local Similarity 72.7%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CASGTASNTTV 11
||:|| || ||
Db 286 CATGTLSNRTV 296

RESULT 14
I48854
gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C;Accession: I48854
R;Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.
Mamm. Genome 5, 726-727, 1994
A;Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.
A;Reference number: I48854; MUID:95178848
A;Accession: I48854
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-459 <RES>
A;Cross-references: EMBL:X76401; NID:g433830; PIDN:CAA53981.1; PID:g433831
C;Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
F;151-188/Domain: NGF receptor repeat homology <NGF>

Query Match 48.2%; Score 41; DB 2; Length 459;
Best Local Similarity 57.1%; Pred. No. 44;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAAD 14
|| || |:|| : |
Db 151 CAPGTFSDTTSSTD 164

RESULT 15
B38634
tumor necrosis factor receptor type 2 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: B38634; A40254; S54816
R;Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen,
Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991
A;Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis facto
A;Reference number: A38634; MUID:91187885
A;Accession: B38634
A;Molecule type: mRNA
A;Residues: 1-474 <LEW>
A;Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R;Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; J
Mol. Cell. Biol. 11, 3020-3026, 1991
A;Title: Molecular cloning and expression of the type 1 and type 2 murine receptors f
A;Reference number: A40254; MUID:91246168
A;Accession: A40254
A;Molecule type: mRNA
A;Residues: 1-474 <GOO>
A;Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R;Kissonerghis, M.; Fellowes, R.; Feldmann, M.; Chernajovsky, Y.
submitted to the EMBL Data Library, May 1995
A;Description: Characterization of the promoter region of the murine p75-TNF receptor
A;Reference number: S54816
A;Accession: S54816
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-22 <KIS>
A;Cross-references: EMBL:X87128; NID:g809043; PIDN:CAA60618.1; PID:g809044
C;Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C;Keywords: cytokine receptor; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-474/Product: tumor necrosis factor receptor type 2 #status predicted <MAT>
F;40-77/Domain: NGF receptor repeat homology <NG1>
F;79-120/Domain: NGF receptor repeat homology <NG2>
F;166-203/Domain: NGF receptor repeat homology <NG4>

Query Match 48.2%; Score 41; DB 2; Length 474;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CASGTASNTTVAAD 14
|| || |:|| : |
Db 166 CAPGTFSDTTSSTD 179

Search completed: March 26, 2002, 13:37:19
Job time: 53 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:40:42 ; Search time 24.63 Seconds
(without alignments)
25.307 Million cell updates/sec

Title: US-09-709-201-96
Perfect score: 85
Sequence: 1 CASGTASNTTVAADRSN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	76	89.4	392	1	OM1P_CHLPS	Q00087 chlamydia p
2	46	54.1	461	1	TNR2_HUMAN	P20333 homo sapien
3	46	54.1	4753	1	LRP_CAEEL	Q04833 caenorhabdi
4	43	50.6	1462	1	TOP2_PEA	O24308 pisum sativ
5	41	48.2	474	1	TNR2_MOUSE	P25119 mus musculu
6	41	48.2	1237	1	DP3A_CHLMU	Q9pjj7 chlamydia m
7	40	47.1	226	1	C553_PARDE	P29967 paracoccus
8	40	47.1	259	1	TRYP_STRGR	P00775 streptomyce
9	39	45.9	237	1	PRCD_RAT	P28073 rattus norv
10	39	45.9	239	1	PRCD_HUMAN	P28072 homo sapien
11	39	45.9	278	1	YFOL_STRTR	P96051 streptococc
12	39	45.9	527	1	IMA_LYCES	O22478 lycopersico
13	39	45.9	710	1	NASC_BACSU	P42434 bacillus su
14	39	45.9	1181	1	MYSC_DICDI	P42522 dictyosteli
15	39	45.9	1319	1	SOS1_MOUSE	Q62245 mus musculu
16	38	44.7	117	1	VGLJ_HSVSB	P36344 simian herp
17	38	44.7	215	1	PRCD_YEAST	P38624 saccharomyc
18	38	44.7	240	1	BIOD_ERWHE	O06899 erwinia her
19	38	44.7	335	1	PIE1_CAEEL	Q94131 caenorhabdi
20	38	44.7	362	1	YEII_ECOLI	P33020 escherichia
21	38	44.7	485	1	VST2_HEVRH	Q00270 hepatitis e
22	38	44.7	621	1	PLAK_MOUSE	Q02257 mus musculu
23	38	44.7	659	1	VST2_HEVME	Q03500 hepatitis e
24	38	44.7	660	1	VST2_HEVBU	P29326 hepatitis e
25	38	44.7	660	1	VST2_HEVMY	Q04611 hepatitis e
26	38	44.7	660	1	VST2_HEVPA	P33426 hepatitis e
27	38	44.7	705	1	TRDN_RABIT	Q28820 oryctolagus
28	38	44.7	728	1	TRDN_HUMAN	Q13061 homo sapien
29	38	44.7	738	1	PLAK_XENLA	P30998 xenopus lae
30	38	44.7	743	1	PLAK_HUMAN	P14923 homo sapien
31	38	44.7	1310	1	VAC3_HELPY	Q48253 helicobacte
32	38	44.7	1355	1	VG46_HSVI1	Q00104 ictalurid h
33	38	44.7	1461	1	TOP2_CANAL	P87078 candida alb
34	37	43.5	120	1	IAA2_STRGS	P20078 streptomyce
35	37	43.5	239	1	CLDE_HUMAN	O95500 homo sapien
36	37	43.5	239	1	CLDE_MOUSE	Q9z0s3 mus musculu
37	37	43.5	323	1	PENP_PENJA	P00798 penicillium
38	37	43.5	326	1	SS10_YEAST	P13130 saccharomyc
39	37	43.5	349	1	VC22_VARV	P34015 variola vir
40	37	43.5	381	1	OPS1_SCHGR	Q94741 schistocerc
41	37	43.5	396	1	PME_BURSO	P24791 burkholderi
42	37	43.5	478	1	HLYD_PASHA	P16534 pasteurella
43	37	43.5	666	1	MUR2_ENTHR	P39046 enterococcu
44	37	43.5	743	1	OCT1_HUMAN	P14859 homo sapien
45	37	43.5	745	1	OCT1_PIG	Q29076 sus scrofa

ALIGNMENTS

```
RESULT   1
OM1P_CHLPS
ID   OM1P_CHLPS     STANDARD;      PRT;   392 AA.
AC   Q00087;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FPN/PRING;
RX   MEDLINE=94103736; PubMed=8277245;
RA   Storey C., Lusher M., Yates P., Richmond S.;
RT   "Evidence for Chlamydia pneumoniae of non-human origin.";
RL   J. Gen. Microbiol. 139:2621-2626(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61096; CAA43409.1; -.
DR   PIR; A40371; A40371.
DR   PIR; S16137; S16137.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    392       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   392 AA;  42069 MW;  88B3C5D90BBA26DB CRC64;


  Query Match             89.4%;  Score 76;  DB 1;  Length 392;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 ASGTASNTTVAADRSN 17
           ||||||||||||||||
Db      92 ASGTASNTTVAADRSN 107
```

```
RESULT   2
TNR2_HUMAN
ID   TNR2_HUMAN     STANDARD;      PRT;   461 AA.
AC   P20333;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR
DE   BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).
GN   TNFRSF1B OR TNFR2 OR TNFBR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90260639; PubMed=2160731;
RA   Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,
RA   Dower S.K., Cosman D., Goodwin R.G.;
RT   "A receptor for tumor necrosis factor defines an unusual family of
RT   cellular and viral proteins.";
RL   Science 248:1019-1023(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91045991; PubMed=2172983;
RA   Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,
RA   Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;
RT   "A second tumor necrosis factor receptor gene product can shed a
RT   naturally occurring tumor necrosis factor inhibitor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96299745; PubMed=8661109;
RA   Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,
RA   Lepaslier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,
RA   Brodeur G.M.;
RT   "Physical mapping and genomic structure of the human TNFR2 gene.";
RL   Genomics 35:94-100(1996).
RN   [4]
RP   SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=90349572; PubMed=2166946;
RA   Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,
RA   Ringold G.M.;
RT   "Complementary DNA cloning of a receptor for tumor necrosis factor
RT   and demonstration of a shed form of the receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
RN   [5]
RP   SEQUENCE OF 27-31.
RX   MEDLINE=90110215; PubMed=2153136;
RA   Engelmann H., Novick D., Wallach D.;
RT   "Two tumor necrosis factor-binding proteins purified from human
RT   urine. Evidence for immunological cross-reactivity with cell surface
RT   tumor necrosis factor receptors.";
RL   J. Biol. Chem. 265:1531-1536(1990).
RN   [6]
RP   SEQUENCE OF 23-40; 65-69; 136-141; 300-306 AND 346-362.
RX   MEDLINE=91056048; PubMed=2173696;
RA   Loetscher H., Schlaeger E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W.,
RA   Brockhaus M.;
RT   "Purification and partial amino acid sequence analysis of two
RT   distinct tumor necrosis factor receptors from HL60 cells.";
RL   J. Biol. Chem. 265:20131-20138(1990).
RN   [7]
RP   CHARACTERIZATION.
RX   MEDLINE=93016040; PubMed=1328224;
RA   Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M.,
RA   Lipari M.T., Goeddel D.V.;
RT   "Biochemical properties of the 75-kDa tumor necrosis factor receptor.
RT   Characterization of ligand binding, internalization, and receptor
RT   phosphorylation.";
RL   J. Biol. Chem. 267:21172-21178(1992).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.
RX   MEDLINE=99221490; PubMed=10206649;
RA   Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;
RT   "Structural basis for self-association and receptor recognition of
RT   human TRAF2.";
RL   Nature 398:533-538(1999).
CC   -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND
CC       APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW
CC       LEVEL ON THREONINE RESIDUES.
CC   -!- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND
CC       WYETH-AYERST). USED TO TREAT MODERATE TO SERVERE RHEUMATOID
CC       ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING
CC       PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO
CC       TNF-ALPHA AND BLOCKS ITS INTERACTIONS WITH RECEPTORS.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD120b entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd120b.htm".
CC   -!- DATABASE: NAME=Enbrel; NOTE=Clinical information on Enbrel;
CC       WWW="http://www.enbrelinfo.com/".
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M32315; AAA59929.1; -.
DR   EMBL; M35857; AAA63262.1; -.
DR   EMBL; U52165; AAC50622.1; -.
DR   EMBL; U52156; AAC50622.1; JOINED.
DR   EMBL; U52157; AAC50622.1; JOINED.
DR   EMBL; U52158; AAC50622.1; JOINED.
DR   EMBL; U52159; AAC50622.1; JOINED.
DR   EMBL; U52160; AAC50622.1; JOINED.
DR   EMBL; U52161; AAC50622.1; JOINED.
DR   EMBL; U52162; AAC50622.1; JOINED.
DR   EMBL; U52163; AAC50622.1; JOINED.
DR   EMBL; U52164; AAC50622.1; JOINED.
DR   EMBL; M55994; AAA36755.1; -.
DR   PIR; A35356; A35356.
DR   PIR; A36007; A36007.
DR   PIR; A36475; A36475.
DR   PIR; B35010; B35010.
DR   PIR; A23666; A23666.
DR   PDB; 1CA9; 12-APR-99.
DR   MIM; 191191; -.
DR   InterPro; IPR001368; TNFR_c6.
DR   Pfam; PF00020; TNFR_c6; 4.
DR   ProDom; PD000771; TNFR_c6; 1.
DR   SMART; SM00208; TNFR; 4.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS50050; TNFR_NGFR_2; 4.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
KW   Phosphorylation; Pharmaceutical; 3D-structure.
FT   SIGNAL        1     22
FT   CHAIN        23    461       TUMOR NECROSIS FACTOR RECEPTOR 2.
FT   DOMAIN       23    257       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    258    287       POTENTIAL.
FT   DOMAIN      288    461       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       39    201       4 X TNFR-CYS.
FT   REPEAT       39     76       TNFR-CYS 1.
FT   REPEAT       77    118       TNFR-CYS 2.
FT   REPEAT      119    162       TNFR-CYS 3.
FT   REPEAT      163    201       TNFR-CYS 4.
FT   DISULFID     40     53       BY SIMILARITY.
FT   DISULFID     54     67       BY SIMILARITY.
FT   DISULFID     57     75       BY SIMILARITY.
FT   DISULFID     78     93       BY SIMILARITY.
FT   DISULFID     96    110       BY SIMILARITY.
FT   DISULFID    100    118       BY SIMILARITY.
```

```
FT   DISULFID    120    126       BY SIMILARITY.
FT   DISULFID    134    143       BY SIMILARITY.
FT   DISULFID    137    161       BY SIMILARITY.
FT   DISULFID    164    179       BY SIMILARITY.
FT   CARBOHYD    171    171       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    193    193       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    141    141       R -> P (IN REF. 4).
FT   CONFLICT    196    196       R -> M (IN REF. 1 AND 3).
FT   CONFLICT    363    363       A -> T (IN REF. 4).
SQ   SEQUENCE   461 AA;  48316 MW;  603B580ECD67636F CRC64;


  Query Match             54.1%;  Score 46;  DB 1;  Length 461;
  Best Local Similarity   64.3%;  Pred. No. 3;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           || || |||| : |
Db     164 CAPGTFSNTTSSTD 177


RESULT   3
LRP_CAEEL
ID   LRP_CAEEL      STANDARD;      PRT;  4753 AA.
AC   Q04833;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN PRECURSOR (LRP).
GN   LRP OR F29D11.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93281621; PubMed=8506301;
RA   Yochem J., Greenwald I.;
RT   "A gene for a low density lipoprotein receptor-related protein in the
RT   nematode Caenorhabditis elegans.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:4572-4576(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Wilkinson J.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY ACT AS A RECEPTOR FOR THE ENDOCYTOSIS OF
CC       EXTRACELLULAR LIGANDS SUCH AS CHYLOMICRON REMNANTS, PROTEASE-
CC       INHIBITOR COMPLEXES AND VITELLOGENIN.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS 35 LDL-RECEPTOR CLASS A DOMAINS.
CC   -!- SIMILARITY: CONTAINS 17 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M96150; AAA28105.1; -.
DR   EMBL; Z73907; CAA98124.1; -.
DR   HSSP; P07204; 1EGT.
DR   WormPep; F29D11.1; CE05765.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002172; LDL_recept_A.
DR   InterPro; IPR000033; Ldl_rcptor_rep.
DR   Pfam; PF00057; ldl_recept_a; 35.
DR   Pfam; PF00058; ldl_recept_b; 26.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   SMART; SM00179; EGF_CA; 2.
DR   SMART; SM00001; EGF_like; 15.
DR   SMART; SM00192; LDLa; 34.
DR   SMART; SM00135; LY; 30.
DR   PROSITE; PS00010; ASX_HYDROXYL; 6.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS01187; EGF_CA; 3.
DR   PROSITE; PS01209; LDLRA_1; 27.
DR   PROSITE; PS50068; LDLRA_2; 34.
KW   Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein;
KW   Signal; Calcium-binding; EGF-like domain; Coated pits.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   4753       LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
FT                                PROTEIN.
FT   DOMAIN       19   4570       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   4571   4596       POTENTIAL.
FT   DOMAIN     4597   4753       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       51     89       LDL-RECEPTOR CLASS A 1.
FT   DOMAIN       90    133       LDL-RECEPTOR CLASS A 2.
FT   DOMAIN      136    177       LDL-RECEPTOR CLASS A 3.
FT   DOMAIN      180    220       LDL-RECEPTOR CLASS A 4.
FT   DOMAIN      221    259       LDL-RECEPTOR CLASS A 5.
FT   DOMAIN      260    298       LDL-RECEPTOR CLASS A 6.
FT   DOMAIN      299    337       EGF-LIKE 1.
FT   DOMAIN      338    368       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      669    712       EGF-LIKE 3.
FT   DOMAIN      997   1043       EGF-LIKE 4.
FT   DOMAIN     1052   1097       LDL-RECEPTOR CLASS A 7.
FT   DOMAIN     1099   1140       LDL-RECEPTOR CLASS A 8.
FT   DOMAIN     1144   1184       LDL-RECEPTOR CLASS A 9.
FT   DOMAIN     1185   1225       LDL-RECEPTOR CLASS A 10.
FT   DOMAIN     1226   1265       LDL-RECEPTOR CLASS A 11.
FT   DOMAIN     1268   1309       LDL-RECEPTOR CLASS A 12.
FT   DOMAIN     1311   1352       LDL-RECEPTOR CLASS A 13.
FT   DOMAIN     1357   1397       LDL-RECEPTOR CLASS A 14.
FT   DOMAIN     1398   1436       EGF-LIKE 5.
FT   DOMAIN     1437   1476       EGF-LIKE 6.
FT   DOMAIN     1747   1786       EGF-LIKE 7.
FT   DOMAIN     2080   2120       EGF-LIKE 8.
FT   DOMAIN     2396   2439       EGF-LIKE 9.
FT   DOMAIN     2728   2780       EGF-LIKE 10.
FT   DOMAIN     2790   2831       LDL-RECEPTOR CLASS A 15.
FT   DOMAIN     2832   2870       LDL-RECEPTOR CLASS A 16.
FT   DOMAIN     2872   2914       LDL-RECEPTOR CLASS A 17.
FT   DOMAIN     2917   2958       LDL-RECEPTOR CLASS A 18.
FT   DOMAIN     2959   2999       LDL-RECEPTOR CLASS A 19.
FT   DOMAIN     3004   3046       LDL-RECEPTOR CLASS A 20.
FT   DOMAIN     3047   3095       LDL-RECEPTOR CLASS A 21.
FT   DOMAIN     3098   3137       LDL-RECEPTOR CLASS A 22.
FT   DOMAIN     3138   3176       LDL-RECEPTOR CLASS A 23.
FT   DOMAIN     3185   3223       LDL-RECEPTOR CLASS A 24.
FT   DOMAIN     3224   3265       EGF-LIKE 11.
FT   DOMAIN     3266   3306       EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3582   3624       EGF-LIKE 13.
FT   DOMAIN     3625   3668       LDL-RECEPTOR CLASS A 25.
FT   DOMAIN     3669   3707       LDL-RECEPTOR CLASS A 26.
FT   DOMAIN     3707   3748       LDL-RECEPTOR CLASS A 27.
FT   DOMAIN     3751   3790       LDL-RECEPTOR CLASS A 28.
FT   DOMAIN     3791   3832       LDL-RECEPTOR CLASS A 29.
FT   DOMAIN     3831   3873       LDL-RECEPTOR CLASS A 30.
FT   DOMAIN     3876   3914       LDL-RECEPTOR CLASS A 31.
FT   DOMAIN     3915   3953       LDL-RECEPTOR CLASS A 32.
FT   DOMAIN     3957   3997       LDL-RECEPTOR CLASS A 33.
FT   DOMAIN     3998   4042       LDL-RECEPTOR CLASS A 34.
FT   DOMAIN     4047   4085       LDL-RECEPTOR CLASS A 35.
FT   DOMAIN     4088   4131       EGF-LIKE 14.
FT   DOMAIN     4132   4176       EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     4477   4515       EGF-LIKE 16.
FT   DOMAIN     4526   4554       EGF-LIKE 17.
FT   DOMAIN     4653   4658       ENDOCYTOSIS SIGNAL (POTENTIAL).
FT   SITE       4744   4744       CRITICAL FOR ENDOCYTOSIS (BY SIMILARITY).
FT   DISULFID     53     65       BY SIMILARITY.
```

```
FT   DISULFID     60     78       BY SIMILARITY.
FT   DISULFID     72     87       BY SIMILARITY.
FT   DISULFID     92    109       BY SIMILARITY.
FT   DISULFID     99    122       BY SIMILARITY.
FT   DISULFID    116    131       BY SIMILARITY.
FT   DISULFID    138    152       BY SIMILARITY.
FT   DISULFID    146    165       BY SIMILARITY.
FT   DISULFID    159    175       BY SIMILARITY.
FT   DISULFID    182    195       BY SIMILARITY.
FT   DISULFID    189    208       BY SIMILARITY.
FT   DISULFID    202    218       BY SIMILARITY.
FT   DISULFID    223    235       BY SIMILARITY.
FT   DISULFID    230    248       BY SIMILARITY.
FT   DISULFID    242    257       BY SIMILARITY.
FT   DISULFID    262    275       BY SIMILARITY.
FT   DISULFID    269    288       BY SIMILARITY.
FT   DISULFID    282    297       BY SIMILARITY.
FT   DISULFID    302    311       BY SIMILARITY.
FT   DISULFID    307    320       BY SIMILARITY.
FT   DISULFID    322    336       BY SIMILARITY.
FT   DISULFID    342    352       BY SIMILARITY.
FT   DISULFID    348    361       BY SIMILARITY.
FT   DISULFID    363    367       BY SIMILARITY.
FT   DISULFID    673    682       BY SIMILARITY.
FT   DISULFID    678    697       BY SIMILARITY.
FT   DISULFID    699    711       BY SIMILARITY.
FT   DISULFID   1001   1010       BY SIMILARITY.
FT   DISULFID   1006   1026       BY SIMILARITY.
FT   DISULFID   1028   1042       BY SIMILARITY.
FT   DISULFID   1054   1068       BY SIMILARITY.
FT   DISULFID   1063   1081       BY SIMILARITY.
FT   DISULFID   1075   1095       BY SIMILARITY.
FT   DISULFID   1101   1114       BY SIMILARITY.
FT   DISULFID   1108   1127       BY SIMILARITY.
FT   DISULFID   1121   1138       BY SIMILARITY.
FT   DISULFID   1146   1158       BY SIMILARITY.
FT   DISULFID   1153   1171       BY SIMILARITY.
FT   DISULFID   1165   1182       BY SIMILARITY.
FT   DISULFID   1187   1199       BY SIMILARITY.
FT   DISULFID   1194   1212       BY SIMILARITY.
FT   DISULFID   1206   1223       BY SIMILARITY.
FT   DISULFID   1228   1241       BY SIMILARITY.
FT   DISULFID   1235   1254       BY SIMILARITY.
FT   DISULFID   1248   1263       BY SIMILARITY.
FT   DISULFID   1270   1283       BY SIMILARITY.
FT   DISULFID   1277   1296       BY SIMILARITY.
FT   DISULFID   1290   1307       BY SIMILARITY.
FT   DISULFID   1313   1325       BY SIMILARITY.
FT   DISULFID   1320   1338       BY SIMILARITY.
FT   DISULFID   1332   1350       BY SIMILARITY.
FT   DISULFID   1359   1373       BY SIMILARITY.
FT   DISULFID   1366   1386       BY SIMILARITY.
FT   DISULFID   1380   1396       BY SIMILARITY.
FT   DISULFID   1401   1412       BY SIMILARITY.
FT   DISULFID   1408   1421       BY SIMILARITY.
FT   DISULFID   1423   1435       BY SIMILARITY.
FT   DISULFID   1441   1451       BY SIMILARITY.
FT   DISULFID   1447   1460       BY SIMILARITY.
FT   DISULFID   1462   1475       BY SIMILARITY.
FT   DISULFID   1751   1760       BY SIMILARITY.
FT   DISULFID   1756   1770       BY SIMILARITY.
FT   DISULFID   1772   1785       BY SIMILARITY.
FT   DISULFID   2084   2095       BY SIMILARITY.
FT   DISULFID   2091   2105       BY SIMILARITY.
FT   DISULFID   2107   2119       BY SIMILARITY.
FT   DISULFID   2400   2415       BY SIMILARITY.
FT   DISULFID   2411   2426       BY SIMILARITY.
FT   DISULFID   2428   2438       BY SIMILARITY.
FT   DISULFID   2732   2743       BY SIMILARITY.
FT   DISULFID   2739   2759       BY SIMILARITY.
FT   DISULFID   2761   2779       BY SIMILARITY.
FT   DISULFID   2792   2805       BY SIMILARITY.
FT   DISULFID   2800   2818       BY SIMILARITY.
FT   DISULFID   2812   2829       BY SIMILARITY.
FT   DISULFID   2834   2846       BY SIMILARITY.
FT   DISULFID   2841   2859       BY SIMILARITY.
FT   DISULFID   2853   2868       BY SIMILARITY.

  Query Match             54.1%;  Score 46;  DB 1;  Length 4753;
  Best Local Similarity   52.9%;  Pred. No. 33;
  Matches    9;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAADRSN 17
           ||:|   | ||| || :
Db    4056 CANGKCVNGTVACDRKD 4072


RESULT   4
TOP2_PEA
ID   TOP2_PEA       STANDARD;      PRT;  1462 AA.
AC   O24308;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   DNA TOPOISOMERASE II (EC 5.99.1.3).
GN   TOP2 OR TOPII.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Reddy M.K., Nair S., Tewari K.K.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC       MAKES DOUBLE-STRAND BREAKS.
CC   -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC       OF DOUBLE-STRANDED DNA.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC       RELAX ONLY NEGATIVE SUPERCOILS.
CC   -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Y14559; CAA74891.1; -.
DR   InterPro; IPR000947; CBFA_NFYB.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   InterPro; IPR003594; HATPase_c.
DR   Pfam; PF00204; DNA_topoisoII; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 2.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00418; TPI2FAMILY.
DR   PRINTS; PR00615; CCAATSUBUNTA.
DR   PRINTS; PR01158; TOPISMRASEII.
DR   ProDom; PD000616; DNA_topoisoII; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA-binding; ATP-binding.
FT   NP_BIND     149    154       ATP (POTENTIAL).
FT   ACT_SITE    761    761       DNA CLEAVAGE (BY SIMILARITY).
SQ   SEQUENCE   1462 AA;  164205 MW;  D9212C54AE0F8B2E CRC64;
```

```
  Query Match           50.6%;  Score 43;  DB 1;  Length 1462;
  Best Local Similarity  50.0%;  Pred. No. 31;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          |::|:|||| :: |
Db   1400 CSAGSASNTPLSED 1413


RESULT   5
TNR2_MOUSE
ID   TNR2_MOUSE     STANDARD;      PRT;   474 AA.
AC   P25119; P97893;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75).
GN   TNFRSF1B OR TNFR2 OR TNFR-2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91187885; PubMed=1849278;
RA   Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA   Wong G.H., Chen E.Y., Goeddel D.V.;
RT   "Cloning and expression of cDNAs for two distinct murine tumor
RT   necrosis factor receptors demonstrate one receptor is species
RT   specific.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91246168; PubMed=1645445;
RA   Goodwin R.G., Anderson D., Jerzy R., Davis T., Brannan C.I.,
RA   Copeland N.G., Jenkins N.A., Smith C.A.;
RT   "Molecular cloning and expression of the type 1 and type 2 murine
RT   receptors for tumor necrosis factor.";
RL   Mol. Cell. Biol. 11:3020-3026(1991).
RN   [3]
RP   SEQUENCE OF 1-26 FROM N.A.
RC   STRAIN=NOD;
RA   Jacob C.O., Liu J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 1-22 FROM N.A.
RC   TISSUE=Liver;
RA   Kissonerghis M., Fellowes R., Feldmann M., Chernajovsky Y.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR TNF-ALPHA.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M60469; AAA39752.1; -.
DR   EMBL; M59378; AAA40463.1; -.
DR   EMBL; U39488; AAA85021.1; -.
DR   EMBL; X87128; CAA60618.1; -.
DR   PIR; B38634; B38634.
DR   HSSP; P19438; 1NCF.
DR   MGD; MGI:1314883; Tnfrsf1b.
DR   InterPro; IPR001368; TNFR_c6.
DR   Pfam; PF00020; TNFR_c6; 4.
DR   ProDom; PD000771; TNFR_c6; 1.
DR   SMART; SM00208; TNFR; 4.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS50050; TNFR_NGFR_2; 3.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    474       TUMOR NECROSIS FACTOR RECEPTOR 2.
FT   DOMAIN       23    258       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    259    288       POTENTIAL.
FT   DOMAIN      289    474       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       39    203       4 X TNFR-CYS.
FT   REPEAT       39     77       TNFR-CYS 1.
FT   REPEAT       78    119       TNFR-CYS 2.
FT   REPEAT      120    164       TNFR-CYS 3.
FT   REPEAT      165    203       TNFR-CYS 4.
FT   DISULFID     40     54       BY SIMILARITY.
FT   DISULFID     55     68       BY SIMILARITY.
FT   DISULFID     58     76       BY SIMILARITY.
FT   DISULFID     79     94       BY SIMILARITY.
FT   DISULFID     97    111       BY SIMILARITY.
FT   DISULFID    101    119       BY SIMILARITY.
FT   DISULFID    121    127       BY SIMILARITY.
FT   DISULFID    136    145       BY SIMILARITY.
FT   DISULFID    139    163       BY SIMILARITY.
FT   DISULFID    166    181       BY SIMILARITY.
FT   CARBOHYD     69     69       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   474 AA;  50319 MW;  462EAE398C4D6563 CRC64;


  Query Match           48.2%;  Score 41;  DB 1;  Length 474;
  Best Local Similarity  57.1%;  Pred. No. 21;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |:|| : |
Db    166 CAPGTFSDTTSSTD 179


RESULT   6
DP3A_CHLMU
ID   DP3A_CHLMU     STANDARD;      PRT;  1237 AA.
AC   Q9PJJ7;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   DNA POLYMERASE III ALPHA SUBUNIT (EC 2.7.7.7).
GN   DNAE OR TC0832.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC       RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC       THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC       THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC       N PYROPHOSPHATE + DNA(N).
CC   -!- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
CC       CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE
CC       DIMERIZATION TO FORM THE POLIII' COMPLEX. POLIII' ASSOCIATES WITH
CC       THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
CC       AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. DNAE
```

```
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002349; AAF39632.1; -.
DR   TIGR; TC0832; -.
DR   InterPro; IPR003141; PHP_N.
DR   Pfam; PF02231; PHP_N; 1.
DR   SMART; SM00481; POLIIIAc; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication;
KW   Complete proteome.
SQ   SEQUENCE   1237 AA;  139892 MW;  5839D4F98D4CA223 CRC64;


  Query Match             48.2%;  Score 41;  DB 1;  Length 1237;
  Best Local Similarity   61.5%;  Pred. No. 55;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       5 TASNTTVAADRSN 17
           | ||| : ||| |
Db     289 TISNTLIVADRCN 301


RESULT   7
C553_PARDE
ID   C553_PARDE     STANDARD;      PRT;   226 AA.
AC   P29967;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C-553I PRECURSOR (C553I).
GN   CYCB.
OS   Paracoccus denitrificans.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Paracoccus.
OX   NCBI_TaxID=266;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=PD 1235;
RX   MEDLINE=92041583; PubMed=1657873;
RA   Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,
RA   Stouthamer A.H.;
RT   "Isolation, sequencing, and mutagenesis of the gene encoding
RT   cytochrome c553i of Paracoccus denitrificans and characterization of
RT   the mutant strain.";
RL   J. Bacteriol. 173:6971-6979(1991).
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC   -!- INDUCTION: DURING GROWTH ON METHANOL.
CC   -!- PTM: BINDS ONE HEME GROUP PER MOLECULE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M75583; AAA25575.1; -.
DR   PIR; B41378; B41378.
DR   InterPro; IPR000345; CytC_heme_bind.
DR   InterPro; IPR003088; Cyt_CI.
DR   Pfam; PF00034; cytochrome_c; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; FALSE_NEG.
KW   Electron transport; Heme; Signal; Methanol utilization; Periplasmic.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23    226       CYTOCHROME C-553I.
FT   BINDING     125    125       HEME (COVALENT) (BY SIMILARITY).
FT   BINDING     128    128       HEME (COVALENT) (BY SIMILARITY).
FT   METAL       129    129       IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT   METAL       173    173       IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT   DOMAIN       37     44       POLY-ALA.
FT   DOMAIN       60     64       POLY-GLU.
SQ   SEQUENCE   226 AA;  23879 MW;  C1D5DAB03702AEC7 CRC64;


  Query Match             47.1%;  Score 40;  DB 1;  Length 226;
  Best Local Similarity   41.2%;  Pred. No. 14;
  Matches    7;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAADRSN 17
           ||:   : | : ||| |
Db      16 CAASAIAGTALCADRRN 32


RESULT   8
TRYP_STRGR
ID   TRYP_STRGR     STANDARD;      PRT;   259 AA.
AC   P00775;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   TRYPSIN PRECURSOR (EC 3.4.21.4) (SGT).
GN   SPRT.
OS   Streptomyces griseus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1911;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 10137;
RX   MEDLINE=92095977; PubMed=1755852;
RA   Kim J.C., Cha S.H., Jeong S.T., Oh S.K., Byun S.M.;
RT   "Molecular cloning and nucleotide sequence of Streptomyces griseus
RT   trypsin gene.";
RL   Biochem. Biophys. Res. Commun. 181:707-713(1991).
RN   [2]
RP   SEQUENCE OF 37-259.
RX   MEDLINE=75127940; PubMed=804314;
RA   Olafson R.W., Jurasek L., Carpenter M.R., Smillie L.B.;
RT   "Amino acid sequence of Streptomyces griseus trypsin. Cyanogen
RT   bromide fragments and complete sequence.";
RL   Biochemistry 14:1168-1177(1975).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX   MEDLINE=88286735; PubMed=3135412;
RA   Read R.J., James M.N.G.;
RT   "Refined crystal structure of Streptomyces griseus trypsin at 1.7-A
RT   resolution.";
RL   J. Mol. Biol. 200:523-551(1988).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M64471; AAA26820.1; ALT_SEQ.
DR   PIR; A00962; TRSMG.
DR   PIR; JQ1302; JQ1302.
DR   PDB; 1SGT; 16-JUL-88.
DR   MEROPS; S01.101; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
```

```
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Zymogen; Signal; 3D-structure.
FT   SIGNAL       1     32
FT   PROPEP       33    36       ACTIVATION PEPTIDE.
FT   CHAIN        37    259      TRYPSIN.
FT   ACT_SITE     73    73       CHARGE RELAY SYSTEM.
FT   ACT_SITE     118   118      CHARGE RELAY SYSTEM.
FT   ACT_SITE     208   208      CHARGE RELAY SYSTEM.
FT   DISULFID     58    74
FT   DISULFID     177   192
FT   DISULFID     204   233
FT   SITE         202   202      REQUIRED FOR SPECIFICITY.
FT   CONFLICT     95    96       MISSING (IN REF. 2).
FT   STRAND       38    38
FT   TURN         39    39
FT   STRAND       41    42
FT   TURN         45    46
FT   TURN         49    50
FT   STRAND       51    54
FT   TURN         55    57
FT   STRAND       58    64
FT   TURN         65    66
FT   STRAND       67    70
FT   HELIX        72    74
FT   STRAND       79    80
FT   STRAND       85    88
FT   STRAND       92    92
FT   TURN         93    94
FT   TURN         96    97
FT   STRAND       99    108
FT   TURN         110   111
FT   STRAND       120   124
FT   STRAND       134   135
FT   TURN         140   141
FT   STRAND       145   150
FT   TURN         156   157
FT   STRAND       163   163
FT   STRAND       165   172
FT   HELIX        174   184
FT   TURN         187   189
FT   STRAND       190   193
FT   TURN         196   198
FT   STRAND       202   202
FT   TURN         205   206
FT   TURN         208   209
FT   STRAND       211   215
FT   TURN         217   218
FT   STRAND       221   229
FT   TURN         236   237
FT   STRAND       240   244
FT   HELIX        245   257
FT   TURN         258   258
SQ   SEQUENCE   259 AA;  26776 MW;  050233AFF1F64823 CRC64;


  Query Match             47.1%;  Score 40;  DB 1;  Length 259;
  Best Local Similarity   46.2%;  Pred. No. 16;
  Matches    6;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAA 13
           | ||: :||:: |
Db      74 CVSGSGNNTSITA 86


RESULT   9
PRCD_RAT
ID   PRCD_RAT       STANDARD;      PRT;   237 AA.
AC   P28073;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROTEASOME DELTA CHAIN PRECURSOR (EC 3.4.99.46) (MACROPAIN DELTA
DE   CHAIN) (MULTICATALYTIC ENDOPEPTIDASE COMPLEX DELTA CHAIN) (PROTEASOME
DE   SUBUNIT Y) (PROTEASOME CHAIN 5) (FRAGMENT).
GN   PSMB6.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93147071; PubMed=1491007;
RA   Tamura T., Shimbara N., Aki M., Ishida N., Bey F., Scherrer K.,
RA   Tanaka K., Ichihara A.;
RT   "Molecular cloning of cDNAs for rat proteasomes: deduced primary
RT   structures of four other subunits.";
RL   J. Biochem. 112:530-534(1992).
RN   [2]
RP   SEQUENCE OF 33-50.
RX   MEDLINE=90242957; PubMed=2335214;
RA   Lilley K.S., Davison M.D., Rivett A.J.;
RT   "N-terminal sequence similarities between components of the
RT   multicatalytic proteinase complex.";
RL   FEBS Lett. 262:327-329(1990).
CC   -!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
CC       WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
CC       ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
CC       NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
CC       PROTEOLYTIC ACTIVITY.
CC   -!- PATHWAY: INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
CC       PROTEOLYTIC PATHWAY.
CC   -!- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
CC       SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC   -!- SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND
CC       ALSO IN THE NUCLEUS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1B; ALSO KNOWN AS THE
CC       PROTEASOME B-TYPE FAMILY. DELTA SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10754; BAA01586.1; ALT_INIT.
DR   PIR; S09086; S09086.
DR   MEROPS; T01.010; -.
DR   InterPro; IPR001353; Proteasome.
DR   InterPro; IPR000243; Proteasome_B.
DR   Pfam; PF00227; proteasome; 1.
DR   PROSITE; PS00854; PROTEASOME_B; 1.
KW   Proteasome; Hydrolase; Protease; Zymogen.
FT   NON_TER      1      1
FT   PROPEP       <1     32
FT   CHAIN        33     237      PROTEASOME DELTA CHAIN.
SQ   SEQUENCE   237 AA;  25158 MW;  C215AC2A70D9AA5F CRC64;


  Query Match             45.9%;  Score 39;  DB 1;  Length 237;
  Best Local Similarity   50.0%;  Pred. No. 21;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           | ||:|::|   ||
Db      76 CRSGSAADTQAVAD 89


RESULT  10
PRCD_HUMAN
ID   PRCD_HUMAN     STANDARD;      PRT;   239 AA.
```

```
AC   P28072;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROTEASOME DELTA CHAIN PRECURSOR (EC 3.4.99.46) (MACROPAIN DELTA
DE   CHAIN) (MULTICATALYTIC ENDOPEPTIDASE COMPLEX DELTA CHAIN) (PROTEASOME
DE   SUBUNIT Y).
GN   PSMB6 OR LMPY.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94345396; PubMed=8066462;
RA   Akiyama K.-Y., Yokota K.-Y., Kagawa S., Shimbara N., Tamura T.,
RA   Akioka H., Nothwang H.G., Noda C., Tanaka K., Ichihara A.;
RT   "cDNA cloning and interferon gamma down-regulation of proteasomal
RT   subunits X and Y.";
RL   Science 265:1231-1234(1994).
RN   [2]
RP   SEQUENCE OF 60-239 FROM N.A., AND SEQUENCE OF 35-75; 80-110 & 210-233.
RX   MEDLINE=91363412; PubMed=1888762;
RA   DeMartino G.N., Orth K., McCullough M.L., Lee L.W., Munn T.Z.,
RA   Moomaw C.R., Dawson P.A., Slaughter C.A.;
RT   "The primary structures of four subunits of the human,
RT   high-molecular-weight proteinase, macropain (proteasome), are
RT   distinct but homologous.";
RL   Biochim. Biophys. Acta 1079:29-38(1991).
RN   [3]
RP   SEQUENCE OF 35-60.
RX   MEDLINE=90167111; PubMed=2306472;
RA   Lee L.W., Moomaw C.R., Orth K., McGuire M.J., DeMartino G.N.,
RA   Slaughter C.A.;
RT   "Relationships among the subunits of the high molecular weight
RT   proteinase, macropain (proteasome).";
RL   Biochim. Biophys. Acta 1037:178-185(1990).
CC   -!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
CC       WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
CC       ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
CC       NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
CC       PROTEOLYTIC ACTIVITY. MAY CATALYZE BASAL PROCESSING OF
CC       INTRACELLULAR ANTIGENS.
CC   -!- PATHWAY: INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
CC       PROTEOLYTIC PATHWAY.
CC   -!- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
CC       SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC   -!- SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND
CC       ALSO IN THE NUCLEUS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1B; ALSO KNOWN AS THE
CC       PROTEASOME B-TYPE FAMILY. DELTA SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D29012; BAA06098.1; -.
DR   EMBL; X61971; CAA43963.1; -.
DR   PIR; S08188; S08188.
DR   PIR; S17522; S17522.
DR   MEROPS; T01.010; -.
DR   SWISS-2DPAGE; P28072; HUMAN.
DR   MIM; 600307; -.
DR   InterPro; IPR001353; Proteasome.
DR   InterPro; IPR000243; Proteasome_B.
DR   Pfam; PF00227; proteasome; 1.
DR   PRINTS; PR00141; PROTEASOME.
DR   PROSITE; PS00854; PROTEASOME_B; 1.
KW   Proteasome; Hydrolase; Protease; Zymogen.
FT   PROPEP        1     34
FT   CHAIN        35    239       PROTEASOME DELTA CHAIN.
FT   CONFLICT    145    145       G -> V (IN REF. 2).
SQ   SEQUENCE   239 AA;  25315 MW;  66EEB9B6C685830D CRC64;


  Query Match             45.9%;  Score 39;  DB 1;  Length 239;
  Best Local Similarity   50.0%;  Pred. No. 22;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAAD 14
           | ||:|::|   ||
Db      78 CRSGSAADTQAVAD 91


RESULT  11
YFOL_STRTR
ID   YFOL_STRTR     STANDARD;      PRT;   278 AA.
AC   P96051;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPOTHETICAL 29.9 KDA PROTEIN IN FOLD-PBP2B INTERGENIC REGION
DE   (ORF1091).
OS   Streptococcus thermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1308;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SFI6;
RX   MEDLINE=97084576; PubMed=8930919;
RA   Stingele F., Mollet B.;
RT   "Disruption of the gene encoding penicillin-binding protein 2b
RT   (pbp2b) causes altered cell morphology and cease in exopolysaccharide
RT   production in Streptococcus thermophilus Sfi6.";
RL   Mol. Microbiol. 22:357-366(1996).
CC   -!- SIMILARITY: BELONGS TO THE UPF0031 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U58210; AAC44613.1; -.
DR   InterPro; IPR000631; UPF0031.
DR   Pfam; PF01256; UPF0031; 1.
DR   PROSITE; PS01049; UPF0031_1; 1.
DR   PROSITE; PS01050; UPF0031_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   278 AA;  29912 MW;  D392FBCB1BE70CD5 CRC64;


  Query Match             45.9%;  Score 39;  DB 1;  Length 278;
  Best Local Similarity   47.1%;  Pred. No. 25;
  Matches    8;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 CASGTASNTTVAADRSN 17
           | :  |   ||| || |
Db      47 CVNSGAGLVTVATDREN 63


RESULT  12
IMA_LYCES
ID   IMA_LYCES      STANDARD;      PRT;   527 AA.
AC   O22478;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   IMPORTIN ALPHA SUBUNIT (KARYOPHERIN ALPHA SUBUNIT) (KAP ALPHA).
```

```
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunik T., Mizrachy L., Citovsky V., Gafni Y.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
CC       EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
CC       SUBSTRATES TO THE NUCLEAR ENVELOPE. SEEMS TO ACT AS A CYTOSOLIC
CC       RECEPTOR FOR BOTH SIMPLE AND BIPARTITE NLS MOTIFS (BY SIMILARITY).
CC   -!- SUBUNIT: FORMS A COMPLEX WITH IMPORTIN BETA-1 SUBUNIT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
CC   -!- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF017252; AAC23722.1; -.
DR   InterPro; IPR000225; Armadillo.
DR   InterPro; IPR002652; IBB.
DR   Pfam; PF00514; Armadillo_seg; 8.
DR   Pfam; PF01749; IBB; 1.
DR   SMART; SM00185; ARM; 8.
DR   PROSITE; PS50176; ARM_REPEAT; 5.
KW   Transport; Protein transport; Repeat.
FT   DOMAIN       12     51       IBB.
FT   REPEAT      109    151       ARM 1.
FT   REPEAT      152    196       ARM 2.
FT   REPEAT      197    234       ARM 3.
FT   REPEAT      235    279       ARM 4.
FT   REPEAT      280    319       ARM 5.
FT   REPEAT      320    362       ARM 6.
FT   REPEAT      363    403       ARM 7.
FT   REPEAT      403    445       ARM 8.
FT   DOMAIN      446    527       ASP/GLU-RICH (ACIDIC).
SQ   SEQUENCE   527 AA;  58605 MW;  4A3F01691CEF4817 CRC64;


  Query Match           45.9%;  Score 39;  DB 1;  Length 527;
  Best Local Similarity  61.5%;  Pred. No. 49;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAAD 14
          ||||: || |  |
Db    145 ASGTSENTKVVID 157


RESULT  13
NASC_BACSU
ID   NASC_BACSU     STANDARD;      PRT;   710 AA.
AC   P42434;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   ASSIMILATORY NITRATE REDUCTASE CATALYTIC SUBUNIT (EC 1.7.99.4).
GN   NASC OR NARB OR NASBB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=97124189; PubMed=8969502;
RA   Yamane K., Kumano M., Kurita K.;
RT   "The 25 degrees-36 degrees region of the Bacillus subtilis
RT   chromosome: determination of the sequence of a 146 kb segment and
RT   identification of 113 genes.";
RL   Microbiology 142:3047-3056(1996).
RN   [2]
RP   SEQUENCE OF 35-710 FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=95173124; PubMed=7868621;
RA   Ogawa K.-I., Akagawa E., Yamane K., Sun Z.-W., Lacelle M., Zuber P.,
RA   Nakano M.M.;
RT   "The nasB operon and nasA gene are required for nitrate/nitrite
RT   assimilation in Bacillus subtilis.";
RL   J. Bacteriol. 177:1409-1413(1995).
CC   -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC       STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC   -!- CATALYTIC ACTIVITY: NITRITE + ACCEPTOR = NITRATE + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN); MAY BIND A 4FE-4S CLUSTER.
CC   -!- PATHWAY: NITRATE ASSIMILATORY PATHWAY.
CC   -!- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
CC       OXIDOREDUCTASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D50453; BAA08965.1; -.
DR   EMBL; D30689; BAA06353.1; -.
DR   EMBL; Z99105; CAB12125.1; -.
DR   HSSP; P07658; 1FDI.
DR   SubtiList; BG11095; nasC.
DR   InterPro; IPR001467; Molybdopterin.
DR   Pfam; PF00384; molybdopterin; 1.
DR   Pfam; PF01568; Molydop_binding; 1.
DR   PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; 1.
DR   PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; 1.
DR   PROSITE; PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
KW   Oxidoreductase; Molybdenum; Nitrate assimilation; Iron-sulfur; 4Fe-4S;
KW   Complete proteome.
FT   METAL        26     26       IRON-SULFUR (4FE-4S) (POTENTIAL).
FT   METAL        29     29       IRON-SULFUR (4FE-4S) (POTENTIAL).
FT   METAL        33     33       IRON-SULFUR (4FE-4S) (POTENTIAL).
FT   METAL        63     63       IRON-SULFUR (4FE-4S) (POTENTIAL).
SQ   SEQUENCE   710 AA;  78575 MW;  625E8864A1552AA2 CRC64;


  Query Match           45.9%;  Score 39;  DB 1;  Length 710;
  Best Local Similarity  57.1%;  Pred. No. 66;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAADR 15
          |: ||:| |  |||
Db    157 AAATAANQTFGADR 170


RESULT  14
MYSC_DICDI
ID   MYSC_DICDI     STANDARD;      PRT;  1181 AA.
AC   P42522;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN IC HEAVY CHAIN.
GN   MYOC OR DMIC.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX2;
RX   MEDLINE=95348228; PubMed=7622596;
RA   Peterson M.D., Novak K.D., Reedy M.C., Ruman J.I., Titus M.A.;
RT   "Molecular genetic analysis of myoC, a Dictyostelium myosin I.";
RL   J. Cell Sci. 108:1093-1103(1995).
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC       ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC   -!- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY
CC       AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L35323; AAC37427.1; -.
DR   HSSP; P08799; 1MND.
DR   DictyDb; DD01090; myoC.
DR   InterPro; IPR001452; SH3.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 2.
DR   Pfam; PF00018; SH3; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   PRINTS; PR00452; SH3DOMAIN.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00242; MYSc; 1.
DR   SMART; SM00326; SH3; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   Myosin; Actin-binding; ATP-binding; Chemotaxis; SH3 domain;
KW   Multigene family.
FT   DOMAIN        1      ?      MYOSIN HEAD-LIKE.
FT   DOMAIN        ?   1181      NON ALPHA-HELICAL, C-TERMINAL DOMAIN.
FT   NP_BIND     109    116      ATP (POTENTIAL).
FT   DOMAIN     1122   1181      SH3.
SQ   SEQUENCE   1181 AA;  132915 MW;  5EB1EE47F0CA8803 CRC64;


  Query Match             45.9%;  Score 39;  DB 1;  Length 1181;
  Best Local Similarity   56.2%;  Pred. No. 1.1e+02;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAADRSN 17
          |||  ::||||  | |
Db    951 ASGLPASTTVAKVRKN 966


RESULT  15
SOS1_MOUSE
ID   SOS1_MOUSE     STANDARD;      PRT;  1319 AA.
AC   Q62245; Q62244;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SON OF SEVENLESS PROTEIN HOMOLOG 1 (SOS-1) (MSOS-1).
GN   SOS1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SWISS; TISSUE=Eye;
RX   MEDLINE=92335328; PubMed=1631150;
RA   Bowtell D., Fu P., Simon M., Senior P.;
RT   "Identification of murine homologues of the Drosophila son of
RT   sevenless gene: potential activators of ras.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:6511-6515(1992).
RN   [2]
RP   STRUCTURE BY NMR OF 415-548.
RX   MEDLINE=97360234; PubMed=9217262;
RA   Koshiba S., Kigawa T., Kim J.-H., Shirouzu M., Bowtell D.,
RA   Yokoyama S.;
RT   "The solution structure of the pleckstrin homology domain of mouse
RT   Son-of-sevenless 1 (mSos1).";
RL   J. Mol. Biol. 269:579-591(1997).
CC   -!- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP (BY
CC       SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MOST EMBRYONIC AND ADULT TISSUES.
CC   -!- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY DOMAIN (DH).
CC   -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RASGEF DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z11574; CAA77662.1; -.
DR   EMBL; Z11578; CAA77665.1; -.
DR   PDB; 1PMS; 15-MAY-97.
DR   MGD; MGI:98354; Sos1.
DR   InterPro; IPR000166; Histone_core.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR000651; RasGEFN.
DR   InterPro; IPR001895; RasGRF_CDC25.
DR   InterPro; IPR000219; RhoGEF.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00617; RasGEF; 1.
DR   Pfam; PF00618; RasGEFN; 1.
DR   Pfam; PF00621; RhoGEF; 1.
DR   SMART; SM00233; PH; 1.
DR   SMART; SM00147; RasGEF; 1.
DR   SMART; SM00229; RasGEFN; 1.
DR   SMART; SM00325; RhoGEF; 1.
DR   PROSITE; PS00720; GDS_CDC25; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   Guanine-nucleotide releasing factor; 3D-structure.
FT   DOMAIN      202    443      DH.
FT   DOMAIN      444    548      PH.
FT   DOMAIN      777    963      RASGEF.
FT   DOMAIN     1244   1247      POLY-PRO.
SQ   SEQUENCE   1319 AA;  150882 MW;  3286088A5BA0A4A6 CRC64;


  Query Match             45.9%;  Score 39;  DB 1;  Length 1319;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAA 13
          ||||:||| | :
Db   1092 ASGTSSNTDVCS 1103


Search completed: March 26, 2002, 13:40:43
Job time: 257 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:40:11 ; Search time 79.01 Seconds
(without alignments)
31.472 Million cell updates/sec

Title: US-09-709-201-96
Perfect score: 85
Sequence: 1 CASGTASNTTVAADRSN 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	76	89.4	392	2	Q99QB0	Q99qb0 chlamydophi
2	54	63.5	356	2	O52924	O52924 chlamydia p
3	54	63.5	390	2	Q9AIJ5	Q9aij5 chlamydia p
4	54	63.5	392	2	Q9AIJ4	Q9aij4 chlamydia p
5	46	54.1	30	4	Q9UIH1	Q9uih1 homo sapien
6	46	54.1	425	4	Q16042	Q16042 homo sapien
7	46	54.1	1518	5	Q21850	Q21850 caenorhabdi
8	45	52.9	390	5	Q9GYJ3	Q9gyj3 caenorhabdi
9	45	52.9	767	12	Q66627	Q66627 equine herp
10	45	52.9	942	10	Q9SUT8	Q9sut8 arabidopsis
11	44	51.8	341	2	Q9X653	Q9x653 streptomyce
12	44	51.8	2911	5	Q9BLV4	Q9blv4 leishmania
13	44	51.8	3040	5	Q9GNY4	Q9gny4 leishmania
14	43	50.6	287	4	Q9H9U0	Q9h9u0 homo sapien
15	43	50.6	376	4	Q9H650	Q9h650 homo sapien
16	43	50.6	425	3	Q12741	Q12741 blastocladi
17	43	50.6	427	4	Q9NT65	Q9nt65 homo sapien
18	43	50.6	4899	5	Q9VR91	Q9vr91 drosophila
19	42	49.4	165	5	Q9N3E7	Q9n3e7 caenorhabdi
20	42	49.4	236	1	Q9HHV1	Q9hhv1 halobacteri
21	42	49.4	390	2	Q9K7M8	Q9k7m8 bacillus ha
22	42	49.4	518	2	O53738	O53738 mycobacteri
23	41.5	48.8	473	2	Q9RK75	Q9rk75 streptomyce
24	41	48.2	119	5	O61034	O61034 trypanosoma
25	41	48.2	175	11	Q9WUL4	Q9wul4 rattus norv
26	41	48.2	217	2	Q9X6X0	Q9x6x0 streptococc
27	41	48.2	338	5	Q9BLA4	Q9bla4 leishmania
28	41	48.2	459	11	Q62327	Q62327 mus musculu
29	41	48.2	482	11	O88734	O88734 mus musculu
30	41	48.2	3257	5	Q9V736	Q9v736 drosophila
31	40	47.1	226	2	Q51672	Q51672 paracoccus
32	40	47.1	259	2	Q54168	Q54168 streptomyce
33	40	47.1	355	10	Q9ATT5	Q9att5 coix lachry
34	40	47.1	490	10	Q9SUY8	Q9suy8 arabidopsis
35	40	47.1	492	10	Q9M044	Q9m044 arabidopsis
36	40	47.1	510	10	Q9M057	Q9m057 arabidopsis
37	40	47.1	760	2	Q47429	Q47429 escherichia
38	39	45.9	126	5	Q9GXD5	Q9gxd5 leishmania
39	39	45.9	166	11	Q9CV50	Q9cv50 mus musculu
40	39	45.9	190	11	Q9CV36	Q9cv36 mus musculu
41	39	45.9	202	5	Q9GZ03	Q9gz03 caenorhabdi
42	39	45.9	202	11	Q60692	Q60692 mus musculu
43	39	45.9	216	2	Q9X6H6	Q9x6h6 streptococc
44	39	45.9	226	3	O43063	O43063 schizosacch
45	39	45.9	231	13	O12992	O12992 lampetra ja

ALIGNMENTS

RESULT 1
Q99QB0
ID Q99QB0 PRELIMINARY; PRT; 392 AA.
AC Q99QB0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN OMPA.
OS Chlamydophila felis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP BAKER, ATCC VR120, AND FP CELLO;
RX MEDLINE=21078580; PubMed=11211261;
RA Bush R.M., Everett K.D.;
RT "Molecular evolution of the Chlamydiaceae.";
RL Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR EMBL; AF269257; AAK00238.1; -.
DR EMBL; AF269258; AAK00239.1; -.
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 392 MAJOR OUTER MEMBRANE PROTEIN.
SQ SEQUENCE 392 AA; 42051 MW; 88B3C09C1FEE26DB CRC64;

Query Match 89.4%; Score 76; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ASGTASNTTVAADRSN 17
||||||||||||||||
Db 92 ASGTASNTTVAADRSN 107

RESULT 2
O52924
ID O52924 PRELIMINARY; PRT; 356 AA.
AC O52924;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER MEMBRANE PROTEIN 1 (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=92-1293, AVIAN SEROVAR D;
RA   Vanrompay D., Cox E., Goddeeris B.M., Volckaert G.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y16562; CAA76286.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane.
FT   NON_TER     356    356
SQ   SEQUENCE   356 AA;  38396 MW;  D51DE06FB46E6F13 CRC64;


  Query Match           63.5%;  Score 54;  DB 2;  Length 356;
  Best Local Similarity   68.8%;  Pred. No. 0.79;
  Matches   11;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAADRSN 17
          |:|||| || | ||:|
Db     92 ATGTASATTTAVDRTN 107


RESULT   3
Q9AIJ5
ID   Q9AIJ5      PRELIMINARY;      PRT;   390 AA.
AC   Q9AIJ5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW JERSEY 1, NJ1;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269266; AAK00247.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     20       POTENTIAL.
FT   CHAIN        21    390       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   390 AA;  42042 MW;  B62858403DBFA4E6 CRC64;


  Query Match           63.5%;  Score 54;  DB 2;  Length 390;
  Best Local Similarity   68.8%;  Pred. No. 0.85;
  Matches   11;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAADRSN 17
          |:|||| || | ||:|
Db     90 ATGTASATTTAVDRTN 105


RESULT   4
Q9AIJ4
ID   Q9AIJ4      PRELIMINARY;      PRT;   392 AA.
AC   Q9AIJ4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TEXAS TURKEY 3, TT3;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269267; AAK00248.1; -.
KW   Signal.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23    392       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   392 AA;  42293 MW;  FC31FC051955246C CRC64;


  Query Match           63.5%;  Score 54;  DB 2;  Length 392;
  Best Local Similarity   68.8%;  Pred. No. 0.86;
  Matches   11;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAADRSN 17
          |:|||| || | ||:|
Db     92 ATGTASATTTAVDRTN 107


RESULT   5
Q9UIH1
ID   Q9UIH1      PRELIMINARY;      PRT;   30 AA.
AC   Q9UIH1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR 2 (FRAGMENT).
GN   TNFR2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Komata T., Tsuchiya N., Matsushita M., Tokunaga K.;
RT   "New poliymorphism within the extracellular region of TNFR2.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB030950; BAA89053.1; -.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3183 MW;  942C00239B909DF5 CRC64;


  Query Match           54.1%;  Score 46;  DB 4;  Length 30;
  Best Local Similarity   64.3%;  Pred. No. 1.5;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db     11 CAPGTFSNTTSSTD 24


RESULT   6
Q16042
ID   Q16042      PRELIMINARY;      PRT;   425 AA.
AC   Q16042;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91370690; PubMed=1966549;
RA   Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,
RA   Brockhaus M., Lesslauer W.;
RT   "Two human TNF receptors have similar extracellular, but distinct
RT   intracellular, domain sequences.";
RL   Cytokine 2:231-237(1990).
DR   EMBL; S63368; AAB19824.2; -.
DR   HSSP; P25942; 1CDF.
DR   InterPro; IPR001368; TNFR_c6.
DR   Pfam; PF00020; TNFR_c6; 4.
DR   ProDom; PD000771; TNFR_c6; 1.
DR   SMART; SM00208; TNFR; 4.
DR   PROSITE; PS00652; TNFR_NGFR_1; 1.
DR   PROSITE; PS50050; TNFR_NGFR_2; 3.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   425 AA;  44608 MW;  1B24A97E3AD4CF9F CRC64;


  Query Match             54.1%;  Score 46;  DB 4;  Length 425;
  Best Local Similarity   64.3%;  Pred. No. 17;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          || || |||| : |
Db    128 CAPGTFSNTTSSTD 141


RESULT   7
Q21850
ID   Q21850      PRELIMINARY;      PRT;  1518 AA.
AC   Q21850;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SIMILAR TO LDL RECEPTOR-RELATED PROTEIN. NCBI GI: 1086710.
GN   R08E3.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Hawkins J.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U41027; AAC48194.1; -.
DR   InterPro; IPR000561; EGF-like.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_4.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
SQ   SEQUENCE   1518 AA;  165773 MW;  35225BE70B7836C0 CRC64;


  Query Match             54.1%;  Score 46;  DB 5;  Length 1518;
  Best Local Similarity   52.9%;  Pred. No. 53;
  Matches    9;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAADRSN 17
          |  |  ||  :| ||||
Db   1275 CVYGLQSNIKIAIDRSN 1291


RESULT   8
Q9GYJ3
ID   Q9GYJ3      PRELIMINARY;      PRT;  390 AA.
AC   Q9GYJ3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C46H11.8 PROTEIN.
GN   C46H11.8.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Miller N., Bradshaw H., Wamsley P.;
RT   "The sequence of C. elegans cosmid C46H11.";
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U88314; AAF99887.1; -.
DR   InterPro; IPR002526; DUF18.
DR   InterPro; IPR003582; ShKT.
DR   Pfam; PF01549; DUF18; 1.
DR   SMART; SM00254; ShKT; 5.
SQ   SEQUENCE   390 AA;  40714 MW;  6158969B404CB161 CRC64;


  Query Match             52.9%;  Score 45;  DB 5;  Length 390;
  Best Local Similarity   56.2%;  Pred. No. 22;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ASGTASNTTVAADRSN 17
          :| |:|::|  |||||
Db    182 SSTTSSSSTTCADRSN 197


RESULT   9
Q66627
ID   Q66627      PRELIMINARY;      PRT;  757 AA.
AC   Q66627;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ORF 24.
OS   Equine herpesvirus type 2 (strain 86/87).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae.
OX   NCBI_TaxID=82831;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95302501; PubMed=7783207;
```

```
RA   Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;
RT   "The DNA sequence of equine herpesvirus 2.";
RL   J. Mol. Biol. 249:520-528(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Telford E.A.R.;
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U20824; AAC13811.1; -.
SQ   SEQUENCE   767 AA;  84112 MW;  6DB41AA51B5B8EDA CRC64;


  Query Match             52.9%;  Score 45;  DB 12;  Length 767;
  Best Local Similarity   57.1%;  Pred. No. 41;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAAD 14
          ||||||  |  :: |
Db    511 CASGTAINMNISGD 524


RESULT  10
Q9SUT8
ID   Q9SUT8      PRELIMINARY;      PRT;   942 AA.
AC   Q9SUT8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   RESPIRATORY BURST OXIDASE HOMOLOG F-LIKE PROTEIN.
GN   F8L21.20 OR AT4G11230.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA   Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL096882; CAB51407.1; -.
DR   EMBL; AL161531; CAB81224.1; -.
DR   InterPro; IPR002916; Ferric_reduct.
DR   InterPro; IPR000778; GP91PhoX.
DR   Pfam; PF01794; Ferric_reduct; 1.
DR   PRINTS; PR00466; GP91PHOX.
SQ   SEQUENCE   942 AA;  107211 MW;  70672D6E2A9BCCD3 CRC64;


  Query Match             52.9%;  Score 45;  DB 10;  Length 942;
  Best Local Similarity   43.8%;  Pred. No. 49;
  Matches    7;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAADRS 16
          | ||: ||: :::| |
Db    769 CISGSCSNSNISSDHS 784


RESULT  11
Q9X653
ID   Q9X653      PRELIMINARY;      PRT;   341 AA.
AC   Q9X653;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   NAPG OXIDOREDUCTASE.
GN   NAPG!
OS   Streptomyces collinus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=42684;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TU 1892;
RX   MEDLINE=99203506; PubMed=10103039;
RA   Chen S., von Bamberg D., Hale V., Breuer M., Hardt B., M ller R.,
RA   Floss H.G., Reynolds K.A., Leistner E.;
RT   "Biosynthesis of ansatrienin (mycotrienin) and naphthomycin.
RT   Identification and analysis of two separate biosynthetic gene clusters
RT   in Streptomyces collinus Tu 1892.";
RL   Eur. J. Biochem. 261:98-107(1999).
DR   EMBL; AF131877; AAD31829.1; -.
DR   InterPro; IPR000683; GFO_IDH_MocA.
DR   InterPro; IPR002965; P_rich_extensn.
DR   Pfam; PF01408; GFO_IDH_MocA; 1.
DR   PRINTS; PR01217; PRICHEXTENSN.
SQ   SEQUENCE   341 AA;  34598 MW;  8809EE179285291E CRC64;


  Query Match             51.8%;  Score 44;  DB 2;  Length 341;
  Best Local Similarity   50.0%;  Pred. No. 28;
  Matches    8;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAADRS 16
          | ||| ::|  ||:|:
Db     97 CLSGTEADTLAAAERA 112


RESULT  12
Q9BLV4
ID   Q9BLV4      PRELIMINARY;      PRT;  2911 AA.
AC   Q9BLV4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   POSSIBLE HIGH MOLECULAR MASS NUCLEAR ANTIGEN (FRAGMENT).
GN   L2230!01.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Zimmermann W., Wambutt R., Ivens A.C., Quail M., Rajandream M.A.,
RA   Barrell B.G.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RX   MEDLINE=98146435; PubMed=9477341;
RA   Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA   Smith D.F.;
RT   "A physical map of the Leishmania major Friedlin genome.";
RL   Genome Res. 8:135-145(1998).
DR   EMBL; AL513062; CAC24680.1; -.
FT   NON_TER    2911   2911
SQ   SEQUENCE   2911 AA;  305079 MW;  45CD9DC05BC57091 CRC64;


  Query Match             51.8%;  Score 44;  DB 5;  Length 2911;
  Best Local Similarity   60.0%;  Pred. No. 2e+02;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 SGTASNTTVAADRSN 17
```

```
          :|:||:|  || |||
Db    939 TGSASSTAAAAGRSN 953


RESULT  13
Q9GNY4
ID   Q9GNY4      PRELIMINARY;      PRT;  3040 AA.
AC   Q9GNY4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   POSSIBLE KINETOPLAST-ASSOCIATED PROTEIN (FRAGMENT).
GN   L5852.01.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Reinhardt R., Klages S., Beck A., Ivens A.C., Murphy L., Quail M.,
RA   Rajandream M.A., Barrell B.G.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RX   MEDLINE=98146435; PubMed=9477341;
RA   Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA   Smith D.F.;
RT   "A physical map of the Leishmania major Friedlin genome.";
RL   Genome Res. 8:135-145(1998).
DR   EMBL; AL499614; CAC18858.1; -.
FT   NON_TER       1      1
FT   NON_TER    3040   3040
SQ   SEQUENCE   3040 AA;  323433 MW;  8F2AD28D3C4021ED CRC64;


  Query Match           51.8%;  Score 44;  DB 5;  Length 3040;
  Best Local Similarity   60.0%;  Pred. No. 2e+02;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 SGTASNTTVAADRSN 17
          :|:||:|  || |||
Db    216 TGSASSTAAAAGRSN 230


RESULT  14
Q9H9U0
ID   Q9H9U0      PRELIMINARY;      PRT;  287 AA.
AC   Q9H9U0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CDNA FLJ12550 FIS, CLONE NT2RM4000698.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK022612; BAB14131.1; -.
SQ   SEQUENCE   287 AA;  31637 MW;  66A71B40BFC53F77 CRC64;


  Query Match           50.6%;  Score 43;  DB 4;  Length 287;
  Best Local Similarity   50.0%;  Pred. No. 35;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAADRS 16
          || ||:|  :  |||:
Db    207 CAPGTSSQFSAGADRA 222


RESULT  15
Q9H650
ID   Q9H650      PRELIMINARY;      PRT;  376 AA.
AC   Q9H650;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CDNA: FLJ22607 FIS, CLONE HSI04846.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HUMAN SMALL INTESTINE;
RA   Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA   Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA   Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK026260; BAB15416.1; -.
DR   InterPro; IPR003323; OTU.
DR   Pfam; PF02338; OTU; 1.
DR   PROSITE; PS50802; OTU; 1.
SQ   SEQUENCE   376 AA;  42440 MW;  911A1020341BC66E CRC64;


  Query Match           50.6%;  Score 43;  DB 4;  Length 376;
  Best Local Similarity   50.0%;  Pred. No. 44;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CASGTASNTTVAADRS 16
          || ||:|  :  |||:
Db    296 CAPGTSSQFSAGADRA 311


Search completed: March 26, 2002, 13:40:13
Job time: 227 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:38:47 ; Search time 81.51 Seconds
(without alignments)
12.723 Million cell updates/sec

Title: US-09-709-201-100
Perfect score: 77
Sequence: 1 CIGLAGTDFANQRP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	77	100.0	14	20	AAW95327	Peptide fragment o
2	41	53.2	112	20	AAY48247	Human prostate can
3	40	51.9	98	11	AAR06602	Cytochrome C-552.
4	40	51.9	1096	21	AAY92323	Human alpha-2-delt
5	39	50.6	16	7	AAP61750	Plasmid pSYC720 en
6	39	50.6	99	16	AAR85208	Mouse ALK-6 extrac
7	39	50.6	126	17	AAR95230	Chick soluble BMP
8	39	50.6	201	17	AAR95232	Chick incomplete B
9	39	50.6	320	21	AAG44529	Arabidopsis thalia
10	39	50.6	348	21	AAG44528	Arabidopsis thalia
11	39	50.6	351	21	AAG44527	Arabidopsis thalia
12	39	50.6	355	21	AAY53884	A suppressor of cy
13	39	50.6	462	19	AAW73012	Cobra venom protea
14	39	50.6	496	22	AAE06222	Booroola sheep mut
15	39	50.6	502	15	AAR55374	Mouse Activin rece
16	39	50.6	502	16	AAR70238	Bone morphogenic p
17	39	50.6	502	16	AAR85209	Mouse ALK-6. Mus
18	39	50.6	502	17	AAR95226	Chick BMP type I r
19	39	50.6	502	17	AAR96202	Bone morphogenetic
20	39	50.6	502	20	AAY33307	Human mALK-6 clone
21	39	50.6	502	20	AAW86249	Mouse BMP receptor
22	39	50.6	502	22	AAE06221	Wild-type sheep BM
23	39	50.6	502	22	AAE06225	Human BMP1B recept
24	39	50.6	521	19	AAW73010	Cobra venom protea
25	39	50.6	592	19	AAW73011	Cobra venom protea
26	38	49.4	277	21	AAG10705	Arabidopsis thalia
27	38	49.4	290	21	AAG34779	Arabidopsis thalia
28	38	49.4	352	21	AAG10704	Arabidopsis thalia
29	38	49.4	356	21	AAG10703	Arabidopsis thalia
30	38	49.4	356	22	AAE01908	Arabidopsis thalia
31	38	49.4	365	21	AAG34778	Arabidopsis thalia
32	38	49.4	369	21	AAG34777	Arabidopsis thalia
33	37.5	48.7	288	18	AAW21776	Protein encoded by
34	37	48.1	135	18	AAW27804	Staphylococcus aur
35	36.5	47.4	319	10	AAP91951	Polypeptide with l
36	36.5	47.4	319	17	AAW09624	Pseudomonas glumae
37	36.5	47.4	319	17	AAW09625	Pseudomonas glumae
38	36.5	47.4	319	17	AAR88018	Mature Pseudomonas
39	36.5	47.4	319	17	AAR88010	Mature Pseudomonas
40	36.5	47.4	319	17	AAR88011	Mature Pseudomonas
41	36.5	47.4	319	17	AAR88012	Mature Pseudomonas
42	36.5	47.4	319	17	AAR88013	Mature Pseudomonas
43	36.5	47.4	319	17	AAR88014	Mature Pseudomonas
44	36.5	47.4	319	17	AAR88015	Mature Pseudomonas
45	36.5	47.4	319	17	AAR88016	Mature Pseudomonas

ALIGNMENTS

RESULT 1
AAW95327
ID AAW95327 standard; Protein; 14 AA.
XX
AC AAW95327;
XX
DT 15-MAR-1999 (first entry)
XX
DE Peptide fragment of C. psitacci CPS160-172.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia psitacci.
XX
PN WO9850074-A2.
XX
PD 12-NOV-1998.
XX
PF 06-MAY-1998; 98WO-US09237.
XX
PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
PR 18-FEB-1998; 98US-0025176.
XX

PA (UYVA-) UNIV VANDERBILT.
XX
PI Mitchell WM, Stratton CW;
XX
DR WPI; 1999-059653/05.
XX
PT Composition with two agents effective against different stages of
PT chlamydial life cycle - comprises agent targetted against cryptic
PT phase, against elementary body phase, against replicating phase,
PT probenicid and antiporphyric
XX
PS Claim 4; Fig 4; 138pp; English.
XX
CC The invention relates to the diagnosis and management of infections by
CC Chlamydia species. The invention provides a composition that comprises
CC at least two agents, where each of the agents is effective against a
CC different phase of the chlamydial life cycle. The agents are selected
CC from: (a) agents targetted against cryptic phase of chlamydial life
CC cycle; (b) agents targetted against elementary body phase of chlamydial
CC life cycle; (c) agents targetted against replicating phase of chlamydial
CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC is used to elicit a protective immune response to Chlamydia infection in
CC an animal or human and is applied until the animal or human tests
CC negative for Chlamydia infection. It is also used to treat biological
CC material infected with Chlamydia. Diagnostic kits for antibody assays
CC against recombinant major outer membrane protein (MOMP), and for DNA
CC amplification assays for chlamydial genes, are used to diagnose disease,
CC e.g. autoimmune disease, an inflammatory disease or a disease that
CC occurs in an immuno-compromised individual, associated with Chlamydia
CC infection. The kits are used to detect chlamydial elementary bodies in a
CC sample. They are also used to monitor and/or modify the course of therapy
CC in a patient. The treatment reduces the acellular load of infectious
CC Ebstein Barr virus. The method is also used to treat porphyria, by
CC reducing the number of elementary bodies and applying a drug, e.g.
CC cimetidine, and antioxidants, to reduce the adverse effects associated
CC with porphyria. Sequences AAW95324 to AAW95327 represent peptides
CC employed for the construction of peptide based ELISAs with species
CC specificity for variable domain 1 (VD1).
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 77; DB 20; Length 14;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CIGLAGTDFANQRP 14
||||||||||||||
Db 1 ciglagtdfanqrp 14

RESULT 2
AAY48247
ID AAY48247 standard; Protein; 112 AA.
XX
AC AAY48247;
XX
DT 08-DEC-1999 (first entry)
XX
DE Human prostate cancer-associated protein 33.
XX
KW Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
KW gene therapy; tissue specificity human.
XX
OS Homo sapiens.
XX
PN DE19811193-A1.
XX
PD 16-SEP-1999.
XX
PF 10-MAR-1998; 98DE-1011193.
XX
PR 10-MAR-1998; 98DE-1011193.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
XX
DR WPI; 1999-519628/44.
DR N-PSDB; AAZ33449.
XX
PT New nucleic acid expressed at high level in prostatic tumor tissue and
PT encoded polypeptides, useful for treating cancer and screening for
PT therapeutic agents -
XX
PS Claim 22; 127; 166pp; German.
XX
CC This invention describes novel nucleic acid sequences (A) that are
CC expressed at high level in prostatic tumor tissue and encode gene
CC products or their fragments. The products of the invention have
CC antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
CC identifying agents for treatment of prostatic cancer and (ii) for
CC therapy of prostate cancer, optionally where expressed by gene therapy
CC methods. (A) is also used to isolate full-length genes (for gene therapy)
CC and for recombinant production of (I), which can be used to raise
CC specific antibodies. (A) are identified by assembly of ESTs (expressed
CC sequence tags) before they are analyzed for expression pattern (tissue
CC specificity). This approach eliminates many of the false results, as
CC regards tissue specificity, associated with known methods that use
CC single (usually short) ESTs. AAY48215-Y48303 represent protein fragments
CC encoded by the expressed sequence tags described in the method of the
CC invention.
XX
SQ Sequence 112 AA;

Query Match 53.2%; Score 41; DB 20; Length 112;
Best Local Similarity 50.0%; Pred. No. 2.9;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CIGLAGTDFANQRP 14
|:|:||| ||
Db 15 cmgvagtalgrarp 28

RESULT 3
AAR06602
ID AAR06602 standard; protein; 98 AA.
XX
AC AAR06602;
XX
DT 11-JAN-1991 (first entry)
XX
DE Cytochrome C-552.
XX
KW Cytochrome C-552; recombinant expression systems; myocardial;
KW infarction; cerebrovacular disease;
XX
OS Hydrogenobacter thermophilus.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label=signal_peptide
XX
PN EP385451-A.
XX
PD 05-SEP-1990.
XX
PF 28-FEB-1990; 90EP-0103928.
XX
PR 28-FEB-1989; 89JP-0047427.
XX
PA (SUNR) SUNTORY LTD.
XX
PI Kodama T, Igarashi Y;
XX

DR WPI; 1990-269480/36.
DR N-PSDB; AAQ05835.
XX
PT New DNA sequence for cytochrome C-552 of Hydrogenobacter
PT thermophilus - and heterologous recombinant expression systems,
PT useful e.g. for treating myocardial infarction and reducing
PT mutagenicity of foods
XX
PS Claim 1; Page 10; 25pp; English.
XX
CC The amino acid sequence of C-552 has been reported and from this
CC information oligonucleotides were synthesised (See AAQ06575-76).
CC The probes were used to isolate a 2.5 kb fragment from a
CC H. thermophilus digest. The C-552 gene was further localised
CC to a 1.1 kb fragment and this was subcloned in pUC14, forming
CC pUHTC 135, and then sequenced. The sequence was inserted into a
CC yeast expression vector, and the recombinants used to transform
CC the yeast strain XS-39-2Bdelta CYC1 (deficient in iso-1
CC cytochrome C so unable to use lactic acid). The gene may also be
CC incorporated into E. coli.
CC The gene product is expected to be useful in treatment of
CC myocardial infarction and cerebrovascular diseases, and, compared
CC with equine cytochrome C, has lower mol. wt. and better stability.
CC It can also be used to reduce mutagenicity of food prods.
XX
SQ Sequence 98 AA;

Query Match 51.9%; Score 40; DB 11; Length 98;
Best Local Similarity 58.3%; Pred. No. 3.9;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IGLAGTDFANQR 13
:|||| |||::
Db 10 vglagitfaneq 21

RESULT 4
AAY92323
ID AAY92323 standard; Protein; 1096 AA.
XX
AC AAY92323;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human alpha-2-delta-D polypeptide from splice variant 1.
XX
KW alpha-2-delta-D; calcium channel; 12p13.3; gabapentin; cytostatic;
KW anticonvulsant; antimigrane; antiparkinsonian; antidepressant;
KW splice variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 310
FT /note= "encoded by RTT"
XX
PN WO200020450-A2.
XX
PD 13-APR-2000.
XX
PF 07-OCT-1999; 99WO-US23519.
XX
PR 07-OCT-1998; 98US-0103322.
PR 30-OCT-1998; 98US-0106473.
PR 29-DEC-1998; 98US-0114088.
XX
PA (WARN) WARNER LAMBERT CO.
XX
PI Johns MA, Moldover B, Offord JD;
XX
DR WPI; 2000-303744/26.
DR N-PSDB; AAA09278.
XX
PT New human nucleic acids encoding the alpha2delta-C and alpha2delta-D
PT proteins, useful in the treatment of epilepsy, migraine, chronic pain,
PT anxiety, multiple sclerosis or cancer
XX
PS Example 3; Page 84; 88pp; English.
XX
CC The alpha-2-delta-D gene encodes a calcium channel subunit polypeptide.
CC The gene has been mapped to chromosome 12p13.1. This gene and the related
CC alpha-2-delta-C and -B genes are useful for protecting mammalian cells
CC from abnormal calcium flux by introducing expression vectors containing
CC the respective gene into mammalian cells. The antisense genes are also
CC useful for treating or preventing epilepsy. The alpha-delta-2-A protein
CC is a high-affinity binding target of the anti-convulsant drug gabapentin.
CC Therefore, alpha-delta-2 proteins may also be targeted to treat
CC seizure-related syndromes, migraine, ataxia, vestibular defects, chronic
CC pain, sleep interference, anxiety, amyotropic lateral sclerosis (ALS),
CC multiple sclerosis, mania, tremor, parkinsonism, substance abuse or
CC addiction syndromes, mood, depression or cancer.
XX
SQ Sequence 1096 AA;

Query Match 51.9%; Score 40; DB 21; Length 1096;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CIGLAGTDFANQRP 14
|: | || || |
Db 1013 clplggtsllnqsp 1026

RESULT 5
AAP61750
ID AAP61750 standard; Protein; 16 AA.
XX
AC AAP61750;
XX
DT 11-JUL-1991 (first entry)
XX
DE Plasmid pSYC720 encoded penP mature protein fragment.
XX
KW Large exopenicillinase; protein secretion.
XX
OS Bacillus licheniformis.
XX
PN ES8506800-A.
XX
PD 16-NOV-1985.
XX
PF 08-MAR-1984; 84ES-0530399.
XX
PR 02-MAR-1984; 84US-0583472.
PR 09-MAR-1983; 83US-0473820.
PR 22-AUG-1989; 89US-0407701.
XX
PA (CETU) CETUS CORP.
XX
PI Shing C;
XX
DR WPI; 1986-057082/09.
DR N-PSDB; AAN60990.
XX
PT Recombinant plasmid vector prodn. - by combining single chain DNA
PT fragments of different lengths.
XX
PS Disclosure; Fig 6; 26pp; English.
XX
CC By removing the Cys residue from the gene product, the membrane
CC binding ability of the secretion signal is interfered with, whilst
CC the transport signal remains active. Thus the secretion of a
CC desired protein may be facilitated at increased levels by
CC attachment of the signal in a suitable host.

```
XX
SQ   Sequence   16 AA;


  Query Match             50.6%;  Score 39;  DB 7;  Length 16;
  Best Local Similarity   46.2%;  Pred. No. 0.8;
  Matches    6;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDFANQR 13
          |: |||: |:| :
Db      1 cvalagsafsnnq 13


RESULT   6
AAR85208
ID   AAR85208 standard; Protein; 99 AA.
XX
AC   AAR85208;
XX
DT   13-FEB-1996  (first entry)
XX
DE   Mouse ALK-6 extracellular domain.
XX
KW   ALK-3; OP1 binding receptor; osteogenic protein 1; morphogenesis;
KW   morphogen; agonist; antagonist.
XX
OS   Mus sp.
XX
PN   WO9530003-A2.
XX
PD   09-NOV-1995.
XX
PF   28-APR-1995;   95WO-US05467.
XX
PR   29-APR-1994;   94US-0236428.
XX
PA   (CREA-) CREATIVE BIOMOLECULES INC.
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Dijke PT,  Heldin C,  Miyazano K,  Sampath KT;
XX
DR   WPI; 1995-393076/50.
DR   N-PSDB; AAT06032.
XX
PT   Identifying osteogenic protein-1 receptor-binding analogue - useful
PT   in the design of morphogen agonists and antagonists for therapeutic,
PT   diagnostic and experimental purposes
XX
PS   Claim 1; Page 73-74; 95pp; English.
XX
CC   The Type-I cell surface receptors ALK-2, ALK-3 and ALK-6 (given in
CC   AAR85206, AAR85207 and AAR85209) have specific binding affinity for
CC   osteogenic protein 1 (OP1) and OP1-related analogues.  The
CC   receptors, and the extracellular domain of ALK-6 (AAR85208), are used
CC   to identify novel morphogen receptor binding analogues useful in
CC   drug design.
XX
SQ   Sequence   99 AA;


  Query Match             50.6%;  Score 39;  DB 16;  Length 99;
  Best Local Similarity   66.7%;  Pred. No. 6.2;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDF 9
          |:|| |:||
Db     48 clglegsdf 56


RESULT   7
AAR95230
ID   AAR95230 standard; Protein; 126 AA.
XX
AC   AAR95230;
XX
DT   31-DEC-1996  (first entry)
XX
DE   Chick soluble BMP type I receptor kinase protein (BRK-2).
XX
KW   BMP type II receptor kinase-3; BRK-3; bone morphogenetic protein;
KW   BMP type I receptor kinase; BRK-2; BMP receptor.
XX
OS   Gallus sp.
XX
PN   WO9614579-A1.
XX
PD   17-MAY-1996.
XX
PF   30-OCT-1995;   95WO-US14027.
XX
PR   05-JUN-1995;   95US-0462467.
PR   04-NOV-1994;   94US-0334178.
XX
PA   (PROC ) PROCTER & GAMBLE CO.
XX
PI   Rosenbaum JS;
XX
DR   WPI; 1996-251887/25.
DR   N-PSDB; AAT28026.
XX
PT   Assays for bone morphogenetic protein activities - using complex of
PT   BMP type I receptor kinase protein and BMP receptor kinase protein
PT   BRK-3
XX
PS   Claim 9; Page 73-74; 101pp; English.
XX
CC   Chick soluble bone morphogenetic protein (BMP) type I receptor
CC   kinase protein-2 (BRK-2) (AAR95230) lacks the regions of the full-
CC   length receptor (AAR95226) not required for BMP binding.  A BMP
CC   receptor kinase protein complex formed of full-length, incomplete
CC   or soluble BMP type I receptor kinase protein and full-length,
CC   incomplete or soluble BMP type II receptor kinase-3 (BRK-3) (see
CC   also AAR95222-29 and AAR95231-34) is useful for screening cpds. for BMP
CC   receptor affinity or for determining the concentration of a BMP
CC   receptor ligand in a clinical sample.  The complex can be expressed
CC   by host cells co-transfected with vectors carrying the appropriate
CC   DNA sequences (see also AAT28018-30).
XX
SQ   Sequence   126 AA;


  Query Match             50.6%;  Score 39;  DB 17;  Length 126;
  Best Local Similarity   66.7%;  Pred. No. 8.2;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDF 9
          |:|| |:||
Db     71 clglegsdf 79


RESULT   8
AAR95232
ID   AAR95232 standard; Protein; 201 AA.
XX
AC   AAR95232;
XX
DT   31-DEC-1996  (first entry)
XX
DE   Chick incomplete BMP type I receptor kinase protein (BRK-2).
XX
KW   BMP type II receptor kinase-3; BRK-3; bone morphogenetic protein;
KW   BMP type I receptor kinase; BRK-2; BMP receptor.
XX
OS   Gallus sp.
XX
PN   WO9614579-A1.
```

XX
PD 17-MAY-1996.
XX
PF 30-OCT-1995; 95WO-US14027.
XX
PR 05-JUN-1995; 95US-0462467.
PR 04-NOV-1994; 94US-0334178.
XX
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Rosenbaum JS;
XX
DR WPI; 1996-251887/25.
DR N-PSDB; AAT28028.
XX
PT Assays for bone morphogenetic protein activities - using complex of
PT BMP type I receptor kinase protein and BMP receptor kinase protein
PT BRK-3
XX
PS Disclosure; Page 77; 101pp; English.
XX
CC Chick incomplete bone morphogenetic protein (BMP) type I receptor
CC kinase protein-1 (BRK-1) (AAR95232) lacks the intracellular kinase
CC domain of the full-length receptor (AAR95226) and is incapable of
CC signal transduction. A BMP receptor kinase protein complex formed
CC of full-length, incomplete or soluble BMP type I receptor kinase
CC protein and full-length, incomplete or soluble BMP type II receptor
CC kinase-3 (BRK-3) (see also AAR95222-31 and AAR95233-34) is useful for
CC screening cpds. for BMP receptor affinity or for determining the
CC concentration of a BMP receptor ligand in a clinical sample. The
CC complex can be expressed by host cells co-transfected with vectors
CC carrying the appropriate DNA sequences (see also AAT28018-30).
XX
SQ Sequence 201 AA;

Query Match 50.6%; Score 39; DB 17; Length 201;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 CIGLAGTDF 9
          |:|| |:||
Db     71 clglegsdf 79
```

RESULT 9
AAG44529
ID AAG44529 standard; Protein; 320 AA.
XX
AC AAG44529;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55790.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.6%; Score 39; DB 21; Length 320;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      6 GTDFANQRP 14
          | || ||||
Db     27 grdfinqrp 35
```

RESULT 10
AAG44528
ID AAG44528 standard; Protein; 348 AA.
XX
AC AAG44528;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55789.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.6%; Score 39; DB 21; Length 348;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      6 GTDFANQRP 14
          | || ||||
Db     55 grdfinqrp 63
```

RESULT 11
AAG44527
ID AAG44527 standard; Protein; 351 AA.
XX
AC AAG44527;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55788.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.6%; Score 39; DB 21; Length 351;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      6 GTDFANQRP 14
          | || ||||
Db     58 grdfinqrp 66
```

RESULT 12
AAY53884
ID AAY53884 standard; Protein; 355 AA.
XX
AC AAY53884;
XX
DT 13-MAR-2000 (first entry)
XX
DE A suppressor of cytokine signalling protein designated HSCOP-4.
XX
KW Human; suppressor of cytokine signalling protein; SOCS protein; HSCOP;
KW cancer; leukaemia; lymphoma; diabetes mellitus; Crohn's disease;
KW immune disorder; AIDS; allergy; atherosclerosis; inflammatory disorder;
KW rheumatoid arthritis; irritable bowel syndrome; multiple sclerosis;
KW ulcerative colitis; neurological disorder; Down's syndrome; amnesia;
KW cerebral neoplasm; Huntingdon's disease; viral infection; adenovirus;
KW acute respiratory disease; toga virus; rubella.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 44
FT /note= "potential phosphorylation site"
FT Modified-site 46
FT /note= "potential phosphorylation site"
FT Modified-site 48
FT /note= "potential phosphorylation site"
FT Modified-site 69
FT /note= "potential phosphorylation site"
FT Modified-site 102
FT /note= "potential phosphorylation site"
FT Modified-site 111
FT /note= "potential glycosylation site"
FT Modified-site 126
FT /note= "potential phosphorylation site"
FT Modified-site 171
FT /note= "potential phosphorylation site"
FT Modified-site 211
FT /note= "potential phosphorylation site"
FT Modified-site 264
FT /note= "potential phosphorylation site"
FT Modified-site 340
FT /note= "potential phosphorylation site"
FT Modified-site 342
FT /note= "potential phosphorylation site"
XX
PN WO9961614-A2.
XX
PD 02-DEC-1999.

```
XX
PF   25-MAY-1999;   99WO-US11497.
XX
PR   28-MAY-1998;   98US-0087104.
PR   17-DEC-1998;   98US-0216006.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Lal P,  Hillman JL,  Gorgone G,  Corley NC,  Patterson C,  Yue H;
PI   Tang YT,  Azimzai Y;
XX
DR   WPI; 2000-072621/06.
DR   N-PSDB; AAZ36828.
XX
PT   New purified polypeptide encoding human suppressor of cytokine
PT   signalling (SOCS) proteins useful for diagnosing, treating or preventing
PT   disorders associated with human SOCS proteins
XX
PS   Claim 1; Page 75-76; 90pp; English.
XX
CC   The present sequence represents a human suppressor of cytokine signalling
CC   (SOCS) protein, designated HSCOP-4. The protein is useful for treating
CC   and/or preventing a disorder associated with decreased expression or
CC   activity of HSCOP. The protein antagonist is useful for treating and/or
CC   preventing a disorder associated with increased expression or activity
CC   of HSCOP. The human SOCS proteins and polynucleotides encoding them are
CC   useful in the diagnosis, treatment and prevention of cancer such as
CC   leukaemia and lymphoma (especially e.g. cancers of the bone, heart and
CC   skin), diabetes mellitus, Crohn's disease, immune disorders e.g. AIDS,
CC   allergies and atherosclerosis, inflammatory disorders e.g. rheumatoid
CC   arthritis, irritable bowel syndrome, multiple sclerosis and ulcerative
CC   colitis, neurological disorders e.g. Down's syndrome, amnesia, cerebral
CC   neoplasms and Huntingdon's disease and infectious diseases such as
CC   those caused by viral infection e.g. adenoviruses (acute respiratory
CC   disease) and toga viruses (rubella) as well as those caused by
CC   bacterial, fungal, parasitic, protozoal and helminthic infections.
XX
SQ   Sequence   355 AA;
```

```
  Query Match             50.6%;  Score 39;  DB 21;  Length 355;
  Best Local Similarity   53.8%;  Pred. No. 26;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDFANQR 13
           |||:| |   |:|
Db     239 cigvaatklqnkr 251
```

```
RESULT  13
AAW73012
ID   AAW73012 standard; Protein; 462 AA.
XX
AC   AAW73012;
XX
DT   02-FEB-1999  (first entry)
XX
DE   Cobra venom protease mocarhagin NMM-3.
XX
KW   Mocarhagin; snake venom; Mozambiquan spitting cobra; protease;
KW   inflammation; myocardial infarction; thrombosis; infection;
KW   metastasis; therapy; NMM-3.
XX
OS   Naja mossambica mossambica.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..61
FT                   /label= Pro_peptide
FT   Protein         62..462
FT                   /label= Mat_protein
XX
PN   WO9846771-A2.
XX
PD   22-OCT-1998.
XX
PF   14-APR-1998;   98WO-US07998.
XX
PR   18-FEB-1998;   98US-0026001.
PR   15-APR-1997;   97US-0843373.
PR   23-JAN-1998;   98US-0012637.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Boodhoo A,  Sako D,  Seehra JS,  Shaw G;
XX
DR   WPI; 1998-568735/48.
DR   N-PSDB; AAV07900.
XX
PT   Isolated mocarhagin cobra venom protease, and nucleic acids encoding
PT   it - used to develop products for treating e.g. myocardial
PT   infarction, thrombosis, bacterial or viral infection, metastatic
PT   conditions or inflammatory disorders
XX
PS   Claim 70; Page 61-63; 97pp; English.
XX
CC   This is the amino acid sequence of mocarhagin NMM-3, a highly
CC   specific metalloproteinase from the venom of the Mozambiquan
CC   spitting cobra.  The invention provides mocarhagin polypeptides
CC   (see AAW73007-13) and polynucleotides (see AAV07895-901) encoding them,
CC   as well as host cells and methods of producing the (especially
CC   mature) polypeptides.  Mocarhagin proteins are capable of cleaving
CC   anionic polypeptide containing sulphated tyrosine residues,
CC   P-selectin glycoprotein (GP) ligand-1 (PSGL-1) and GP1b-alpha
CC   (claimed).  They also inhibit neutrophil/HL60 binding, inhibit
CC   platelet binding to von Willebrand Factor, require Ca2+ and Zn2+
CC   ions for activity and have activity inhibited by excess EDTA or
CC   high concentrations of DFP (claimed).  They can be used to inhibit
CC   selectin-mediated binding and to treat inflammatory disease
CC   (claimed).  In particular, they can be used to treat e.g. myocardial
CC   infarction, vessel restenosis, thrombosis, bacterial or viral
CC   infection, metastatic conditions, inflammatory disorders such as
CC   arthritis, acute respiratory distress syndrome, asthma, emphysema,
CC   delayed type hypersensitivity reaction, systemic lupus
CC   erythematosus, thermal injury such as burns or frostbite,
CC   autoimmune thyroiditis, experimental allergic encephalomyelitis,
CC   multiple sclerosis, multiple organ injury syndrome secondary to
CC   trauma, diabetes, Reynaud's syndrome, neutrophilic dermatosis
CC   (Sweet's syndrome), inflammatory bowel disease, Grave's disease,
CC   glomerulonephritis, gingivitis, periodontitis, haemolytic uremic
CC   syndrome, ulcerative colitis, Crohn's disease, necrotising
CC   enterocolitis, granulocyte transfusion associated syndrome,
CC   cytokine-induced enterocolitis, granulocyte transfusion associated
CC   syndrome, or cytokine-induced toxicity.  Mocarhagin protein may
CC   also be useful in organ transplantation, both to prepare organs for
CC   transplantation and to quell organ transplant rejection, to treat
CC   haemodialysis and leukophoresis patients, or as an inhibitor of P-
CC   or E-selectin-mediated intercellular adhesion.
XX
SQ   Sequence   462 AA;
```

```
  Query Match             50.6%;  Score 39;  DB 19;  Length 462;
  Best Local Similarity   77.8%;  Pred. No. 35;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           || | ||||
Db     369 cialmgtdf 377
```

```
RESULT  14
AAE06222
ID   AAE06222 standard; Protein; 496 AA.
XX
AC   AAE06222;
XX
```

```
DT   25-SEP-2001  (first entry)
XX
DE   Booroola sheep mutant BMP1B receptor.
XX
KW   Bone morphogenetic protein 1B receptor; BMP1B receptor; ovulation;
KW   FecB; fecundity; fertility; Booroola; contraceptive; mutant; mutein;
KW   transforming growth factor-beta family; chromosome 6; sheep.
XX
OS   Ovis aries.
XX
FH   Key             Location/Qualifiers
FT   Domain          127..148
FT                   /label= Membrane_spanning_domain
FT   Domain          203..496
FT                   /label= Intracellular_kinase_domain
FT   Misc-difference 249
FT                   /note= "Wild type Gln is substituted with Arg"
XX
PN   WO200148204-A1.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-NZ00259.
XX
PR   23-DEC-1999;   99NZ-0502058.
XX
PA   (AGRE-) AGRESEARCH LTD.
XX
PI   Wilson TM,  Wu X;
XX
DR   WPI; 2001-441717/47.
DR   N-PSDB; AAD11872.
XX
PT   Novel isolated bone morphogenetic protein 1B receptor polypeptide
PT   useful for modulating the ovulation rate of a female vertebrate -
XX
PS   Claim 12; Page 50-52; 53pp; English.
XX
CC   The invention relates to a bone morphogenetic protein 1B (BMP1B)
CC   receptor polypeptide comprising a sequence which differs from that
CC   of the wild type in that residue 249 is arginine and not glutamine,
CC   or a sequence in which residue 249 is glutamine, but which is otherwise
CC   different from the wild type BMP1B polypeptide sequence and which has
CC   the ability to modulate ovulation in a female mammal. Mutation in
CC   the BMP1B receptor gene is responsible for increased ovulation rate
CC   in sheeps derived from Booroola Merino strain that carry an autosomal
CC   mutation in FecB/Booroola gene. The FecB gene is mapped to chromosome 6.
CC   The BMP1B receptor of the invention and the polynucleotide encoding
CC   it are useful for modulating the ovulation rate of a female vertebrate.
CC   Identification of mutated BMP1B receptor nucleic acid molecule in a
CC   vertebrate, is useful for assessing fecundity in vertebrate such as
CC   humans and other commercially important mammals and birds including
CC   sheep, cattle, horses, goats, deer, pigs, cats, dogs, possums, and
CC   poultry. The polypeptide is useful to raise antibodies and for reducing
CC   unwanted populations of feral vertebrates. The polynucleotide is useful
CC   for identifying sequence variants in individual animals that are
CC   associated with increased ovulation. The present sequence is mutant
CC   BMP1B receptor from Booroola sheep. The BMP1B receptor is a
CC   member of transforming growth factor-beta family.
XX
SQ   Sequence   496 AA;

  Query Match           50.6%;  Score 39;  DB 22;  Length 496;
  Best Local Similarity   66.7%;  Pred. No. 38;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 clglegsdf 79

RESULT  15
```

```
AAR55374
ID   AAR55374 standard; Protein; 502 AA.
XX
AC   AAR55374;
XX
DT   20-JAN-1995  (first entry)
XX
DE   Mouse Activin receptor-like kinase 6 (mALK-6).
XX
KW   serine threonine kinases; activin receptors; Act-R; superfamily;
KW   transforming growth factor; TGF; diagnostics; detection; therapy;
KW   rheumatoid arthritis; glomerular nephritis; fibrosis.
XX
OS   Mus musculus.
XX
PN   WO9411502-A.
XX
PD   26-MAY-1994.
XX
PF   17-NOV-1993;   93WO-GB02367.
XX
PR   17-NOV-1992;   92GB-0024057.
PR   08-MAR-1993;   93GB-0004677.
PR   08-MAR-1993;   93GB-0004680.
PR   28-MAY-1993;   93GB-0011047.
PR   02-JUL-1993;   93GB-0013763.
PR   03-AUG-1993;   93GB-0016099.
PR   15-OCT-1993;   93GB-0021344.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Dijke P,  Franzen P,  Heldin C,  Miyazono K,  Yamashita H;
XX
DR   WPI; 1994-183503/22.
DR   N-PSDB; AAQ66642.
XX
PT   Activin receptor-like kinase(s) with serine/threonine kinase
PT   domains - have activin/TGF beta-type I receptor function and can
PT   be used in diagnosis or therapy of rheumatoid arthritis,
PT   glomerular nephritis, fibrosis, etc.
XX
PS   Claim 3; Page 75-77; 97pp; English.
XX
CC   The inventors have identified a new family of receptor kinases
CC   called activin receptor-like kinases (ALK).  Their discovery was
CC   based on the realisation that receptor serine/threonine kinases
CC   form a new receptor family, which may include the type II receptors
CC   for other proteins in the transforming growth factor(TGF) beta
CC   superfamily. The activin receptor type II sequences from mouse and
CC   the dafI gene product of C.elegans have high sequence similarity
CC   and were used to design degenerate primers to clone related cDNA's
CC   (see AAQ66643-49). Six distinct putative receptor serine/threonine
CC   kinases were identified, called ALK (human ALK proteins are shown
CC   in AAR55366-70, mouse ALK are shown in AAR55371-74). Products of the
CC   invention can be used in therapy, eg. to modulate conditions
CC   associated with activin or TGF beta activity. These conditions
CC   include fibrosis, eg. liver cirrhosis and pulmonary fibrosis, cancer,
CC   rheumatoid arthritis and glomeronephritis.
XX
SQ   Sequence   502 AA;

  Query Match           50.6%;  Score 39;  DB 15;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 39;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 clglegsdf 79


Search completed: March 26, 2002, 13:38:48
```

Job time: 142 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:41:28 ; Search time 37.72 Seconds
(without alignments)
8.352 Million cell updates/sec

Title: US-09-709-201-100
Perfect score: 77
Sequence: 1 CIGLAGTDFANQRP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	51.9	98	6	5459046-6	Patent No. 5459046
2	39	50.6	502	2	US-08-481-337A-8	Sequence 8, Appli
3	39	50.6	502	4	US-09-382-256-18	Sequence 18, Appl
4	39	50.6	502	4	US-09-395-115-18	Sequence 18, Appl
5	39	50.6	502	4	US-08-123-934A-4	Sequence 4, Appli
6	39	50.6	502	4	US-08-334-179A-14	Sequence 14, Appl
7	39	50.6	502	5	PCT-US94-10080-4	Sequence 4, Appli
8	39	50.6	502	5	PCT-US95-05467-8	Sequence 8, Appli
9	36.5	47.4	358	1	US-08-034-650-10	Sequence 10, Appl
10	36.5	47.4	358	1	US-08-449-015-10	Sequence 10, Appl
11	36	46.8	309	2	US-08-849-480A-6	Sequence 6, Appli
12	36	46.8	1580	2	US-08-804-227C-11	Sequence 11, Appl
13	36	46.8	1580	2	US-08-804-198-5	Sequence 5, Appli
14	35	45.5	672	1	US-07-841-651-2	Sequence 2, Appli
15	35	45.5	672	1	US-07-841-651-3	Sequence 3, Appli
16	35	45.5	1612	1	US-08-169-927-2	Sequence 2, Appli
17	34	44.2	176	1	US-08-145-995A-3	Sequence 3, Appli
18	34	44.2	176	2	US-08-451-747-3	Sequence 3, Appli
19	34	44.2	176	3	US-09-134-852-3	Sequence 3, Appli
20	34	44.2	269	4	US-09-028-366-6	Sequence 6, Appli
21	34	44.2	570	2	US-08-484-993B-16	Sequence 16, Appl
22	34	44.2	570	2	US-08-484-158B-16	Sequence 16, Appl
23	34	44.2	570	2	US-08-484-596A-16	Sequence 16, Appl
24	34	44.2	570	2	US-08-480-150A-16	Sequence 16, Appl
25	34	44.2	570	3	US-08-458-731-16	Sequence 16, Appl
26	34	44.2	570	3	US-08-149-223A-16	Sequence 16, Appl
27	34	44.2	591	1	US-08-145-995A-21	Sequence 21, Appl
28	34	44.2	591	2	US-08-451-747-21	Sequence 21, Appl
29	34	44.2	591	3	US-09-134-852-21	Sequence 21, Appl
30	34	44.2	659	4	US-09-189-462-4	Sequence 4, Appli
31	33	42.9	526	3	US-08-504-878A-2	Sequence 2, Appli
32	33	42.9	956	3	US-08-772-270A-8	Sequence 8, Appli
33	33	42.9	1127	4	US-08-937-195-3	Sequence 3, Appli
34	33	42.9	1127	4	US-08-915-152-3	Sequence 3, Appli
35	33	42.9	1127	5	PCT-US96-07627-3	Sequence 3, Appli
36	32	41.6	18	2	US-07-876-941A-29	Sequence 29, Appl
37	32	41.6	18	3	US-08-293-728-20	Sequence 20, Appl
38	32	41.6	18	4	US-09-421-868-20	Sequence 20, Appl
39	32	41.6	157	1	US-08-450-065-2	Sequence 2, Appli
40	32	41.6	157	1	US-08-450-595-2	Sequence 2, Appli
41	32	41.6	235	1	US-08-015-985-13	Sequence 13, Appl
42	32	41.6	252	2	US-08-685-992-33	Sequence 33, Appl
43	32	41.6	252	2	US-09-144-925-33	Sequence 33, Appl
44	32	41.6	255	1	US-08-459-264-4	Sequence 4, Appli
45	32	41.6	255	1	US-08-459-263-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT   1
5459046-6
;Patent No. 5459046
;     APPLICANT: KODAMA, TOHRU;IGARASHI, YASUO
;     TITLE OF INVENTION: CYTOCHROME C GENE DERIVED FROM HYDROGEN
;BACTERIUM
;     NUMBER OF SEQUENCES: 17
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/943,140
;       FILING DATE: 10-SEP-1992
;     PRIOR APPLICATION DATA:486,409
;       APPLICATION NUMBER: 28-FEB-1990
;SEQ ID NO:6:
;       LENGTH: 98
5459046-6


  Query Match             51.9%;  Score 40;  DB 6;  Length 98;
  Best Local Similarity   58.3%;  Pred. No. 2.5;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 IGLAGTDFANQR 13
           :||||  |||::
Db      10 VGLAGITFANEQ 21


RESULT   2
US-08-481-337A-8
; Sequence 8, Application US/08481337A
; Patent No. 5863738
;  GENERAL INFORMATION:
;     APPLICANT:  TEN DIJKE, Peter
;     APPLICANT:  HELDIN, Carl-Henrik
;     APPLICANT:  MIYAZONO, Kohei
;     APPLICANT:  SAMPATH, Kuber T.
;     TITLE OF INVENTION:  Morphogenic Protein-Specific Cell
;     TITLE OF INVENTION:  Surface Receptors and Uses Therefor
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Testa, Hurwitz & Thibeault
;       STREET:  125 High St.
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02110
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,337A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: CRP-097CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-481-337A-8

Query Match 50.6%; Score 39; DB 2; Length 502;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CIGLAGTDF 9
|:|| |:||
Db 71 CLGLEGSDF 79

RESULT 3
US-09-382-256-18
; Sequence 18, Application US/09382256A
; Patent No. 6207814
; GENERAL INFORMATION:
; APPLICANT: MIYAZONO, Kohei
; TEN DIJKE, Peter
; FRANZEN, Petra
; YAMASHITA, Hidetoshi
; HELDIN, Carl-Henrik
; TITLE OF INVENTION: ACTIVIN RECEPTOR LIKE KINASES, PROTEINS
; HAVING SERINE THREONINE KINASE DOMAINS,
; AND THEIR USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/382,256A
; FILING DATE: 24-Aug-1999
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/02367
; FILING DATE: No. 6207814ember 17, 1993
; APPLICATION NUMBER: GB 9224057.1
; FILING DATE: No. 6207814ember 17, 1992
; APPLICATION NUMBER: GB 9304677.9
; FILING DATE: March 8, 1993
; APPLICATION NUMBER: GB 9304680.3
; FILING DATE: March 8, 1993
; APPLICATION NUMBER: 9311047.6
; FILING DATE: May 28, 1993
; APPLICATION NUMBER: 9313763.6
; FILING DATE: July 2, 1993
; APPLICATION NUMBER: 9316099.2
; FILING DATE: August 3, 1993
; APPLICATION NUMBER: 321344.5
; FILING DATE: October 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6207814man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5298.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-382-256-18

Query Match 50.6%; Score 39; DB 4; Length 502;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CIGLAGTDF 9
|:|| |:||
Db 71 CLGLEGSDF 79

RESULT 4
US-09-395-115-18
; Sequence 18, Application US/09395115
; Patent No. 6271365
; GENERAL INFORMATION:
; APPLICANT: Miyazono, Kohei; Dijke, Peter Ten;
; APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik
; TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins
; TITLE OF INVENTION: Having Serine Threonine Kinase Domains And Their Use
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/395,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,265
; FILING DATE: 30-October-1995
; APPLICATION NUMBER: PCT/GB93/02367
; FILING DATE: 17-No. 6271365ember-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9224057.1
; FILING DATE: 17-No. 6271365ember-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9304677.9
; FILING DATE: 8-March-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9304680.3
; FILING DATE: 8-March-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9311047.6
; FILING DATE: 28-May-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9313763.6

```
;       FILING DATE:  2-July-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  9136099.2
;       FILING DATE:  3-August-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  9321344.5
;       FILING DATE:  15-October-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kohlei, Vineet
;       REGISTRATION NUMBER:  37,003
;       REFERENCE/DOCKET NUMBER:  LUD 5298
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 688-9200
;       TELEFAX:  (212) 838-3884
;   INFORMATION FOR SEQ ID NO:  18:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  502 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-395-115-18


  Query Match             50.6%;  Score 39;  DB 4;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 24;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 CLGLEGSDF 79


RESULT   5
US-08-123-934A-4
; Sequence 4, Application US/08123934A
; Patent No. 6291206
;  GENERAL INFORMATION:
;    APPLICANT:  WOZNEY, John
;    APPLICANT:  CELESTE, Anthony J.
;    APPLICANT:  THIES, R. Scott
;    APPLICANT:  YAMAJI, No. 6291206oru
;    TITLE OF INVENTION:  RECEPTOR PROTEINS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute Inc.- Legal Affairs
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,934A
;      FILING DATE:  17-SEP-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LAZAR, Steven R
;      REGISTRATION NUMBER:  32,618
;      REFERENCE/DOCKET NUMBER:  5203
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 876 1170
;      TELEFAX:  617 876 5851
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  502 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-123-934A-4


  Query Match             50.6%;  Score 39;  DB 4;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 24;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 CLGLEGSDF 79


RESULT   6
US-08-334-179A-14
; Sequence 14, Application US/08334179A
; Patent No. 6306622
;  GENERAL INFORMATION:
;    APPLICANT:  ROSENBAUM, JAN S.
;    APPLICANT:  NOHNO, TSUTOMU
;    TITLE OF INVENTION:  cDNA ENCODING A BMP TYPE II RECEPTOR
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THE PROCTER AND GAMBLE COMPANY
;      STREET:  11810 EAST MIAMI RIVER ROAD
;      CITY:  ROSS
;      STATE:  OH
;      COUNTRY:  US
;      ZIP:  45061
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.30, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/334,179A
;      FILING DATE:  04-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CORSTANJE, BRAHM J.
;      REGISTRATION NUMBER:  34,804
;      REFERENCE/DOCKET NUMBER:  5473
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  513-627-2858
;      TELEFAX:  513-627-0260
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  502 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-334-179A-14


  Query Match             50.6%;  Score 39;  DB 4;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 24;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 CLGLEGSDF 79


RESULT   7
PCT-US94-10080-4
; Sequence 4, Application PC/TUS9410080
;  GENERAL INFORMATION:
;    APPLICANT:  GENETICS INSTITUTE, INC.
;    TITLE OF INVENTION:  RECEPTOR PROTEINS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute Inc.- Legal Affairs
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
```

```
;       COUNTRY:  USA
;       ZIP:  02140
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/10080
;       FILING DATE:  HEREWITH
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/123,934
;       FILING DATE:  17-SEP-1993
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  LAZAR, Steven R
;       REGISTRATION NUMBER:  32,618
;       REFERENCE/DOCKET NUMBER:  5203-PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 498-8260
;       TELEFAX:  (617) 876-5851
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  502 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
PCT-US94-10080-4


  Query Match             50.6%;  Score 39;  DB 5;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 24;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 CLGLEGSDF 79


RESULT   8
PCT-US95-05467-8
; Sequence 8, Application PC/TUS9505467
;  GENERAL INFORMATION:
;     APPLICANT:
;     APPLICANT:
;     TITLE OF INVENTION:  MORPHOGENIC PROTEIN-SPECIFIC CELL
;     TITLE OF INVENTION:  SURFACE RECEPTORS AND USES THEREFOR
;     NUMBER OF SEQUENCES:  15
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  PATENT ADMINISTRATOR, TESTA, HURWITZ &
;       ADDRESSEE:  THIBEAULT
;       STREET:  53 STATE STREET
;       CITY:  BOSTON
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/05467
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PITCHER, EDMUND R.
;       REGISTRATION NUMBER:  27,829
;       REFERENCE/DOCKET NUMBER:  CRP-097PC
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 248-7000
;       TELEFAX:  (617) 248-7100
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  502 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
PCT-US95-05467-8


  Query Match             50.6%;  Score 39;  DB 5;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 24;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 CLGLEGSDF 79


RESULT   9
US-08-034-650-10
; Sequence 10, Application US/08034650
; Patent No. 5641671
;  GENERAL INFORMATION:
;     APPLICANT:  BOS, Jannetje W.
;     APPLICANT:  FRENKEN, Leon G.
;     APPLICANT:  VERRIPS, Cornelis T.
;     APPLICANT:  VISSER, Christiaan
;     TITLE OF INVENTION:  PRODUCTION OF ACTIVE PSEUDOMONAS GLUMAE
;     TITLE OF INVENTION:  LIPASE IN HOMOLOGOUS OR HETEROLOGOUS HOSTS
;     NUMBER OF SEQUENCES:  13
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
;       STREET:  1615 L. Street, N.W.
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  USA
;       ZIP:  20036-5601
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/034,650
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/727,235
;       FILING DATE:  03-JUL-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kokulis, Paul N.
;       REGISTRATION NUMBER:  16,773
;       REFERENCE/DOCKET NUMBER:  PNK/5970/91731
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 861-3000
;       TELEFAX:  (202) 822-0944
;       TELEX:  6714627 CUSH
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  358 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-034-650-10


  Query Match             47.4%;  Score 36.5;  DB 1;  Length 358;
  Best Local Similarity   90.0%;  Pred. No. 48;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       3 GLAGTD-FAN 11
           |||||| |||
Db      55 GLAGTDKFAN 64
```

```
RESULT  10
US-08-449-015-10
; Sequence 10, Application US/08449015
; Patent No. 5804409
;  GENERAL INFORMATION:
;    APPLICANT:  BOS, Jannetje W.
;    APPLICANT:  FRENKEN, Leon G.
;    APPLICANT:  VERRIPS, Cornelis T.
;    APPLICANT:  VISSER, Christiaan
;    TITLE OF INVENTION:  PRODUCTION OF ACTIVE PSEUDOMONAS GLUMAE
;    TITLE OF INVENTION:  LIPASE IN HOMOLOGOUS OR HETEROLOGOUS HOSTS
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
;      STREET:  1615 L. Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036-5601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/449,015
;      FILING DATE:  24-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/727,235
;      FILING DATE:  03-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kokulis, Paul N.
;      REGISTRATION NUMBER:  16,773
;      REFERENCE/DOCKET NUMBER:  PNK/5970/91731
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  358 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-449-015-10


  Query Match             47.4%;  Score 36.5;  DB 1;  Length 358;
  Best Local Similarity   90.0%;  Pred. No. 48;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       3 GLAGTD-FAN 11
           |||||| ||| 
Db      55 GLAGTDKFAN 64


RESULT  11
US-08-849-480A-6
; Sequence 6, Application US/08849480A
; Patent No. 5981184
;  GENERAL INFORMATION:
;    APPLICANT:  MELCHERS, Klaus
;    TITLE OF INVENTION:  SCREENING MODEL
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JACOBSON, PRICE, HOLMAN & STERN, PLLC
;      STREET:  400 - 7th Street, N. W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/849,480A
;      FILING DATE:  02-JUN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04711
;      FILING DATE:  30-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P4442970.3
;      FILING DATE:  02-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 19505645.0
;      FILING DATE:  18-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  AISENBERG, Irwin M.
;      REGISTRATION NUMBER:  19,007
;      REFERENCE/DOCKET NUMBER:  8125/P60984US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-6666
;      TELEFAX:  202/393-5350
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  309 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    FRAGMENT TYPE:  C-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Helicobacter pylori
;      STRAIN:  Helicobacter pylori 69A
;      INDIVIDUAL ISOLATE:  Clinical isolate 69A
;    IMMEDIATE SOURCE:
;      LIBRARY:  Helicobacter 69A - gene library in vector
;      LIBRARY:  pRH160
;      CLONE:  pRH948
US-08-849-480A-6


  Query Match             46.8%;  Score 36;  DB 2;  Length 309;
  Best Local Similarity   77.8%;  Pred. No. 50;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GLAGTDFAN 11
           |||| | ||  
Db      52 GLAGADLAN 60


RESULT  12
US-08-804-227C-11
; Sequence 11, Application US/08804227C
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1580 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-804-227C-11


  Query Match             46.8%;  Score 36;  DB 2;  Length 1580;
  Best Local Similarity   63.6%;  Pred. No. 3.1e+02;
  Matches    7;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 LAGTDFANQRP 14
           ||| |||   |
Db     752 LAGVDFAGHSP 762


RESULT 13
US-08-804-198-5
; Sequence 5, Application US/08804198
; Patent No. 5945320
;  GENERAL INFORMATION:
;    APPLICANT:  Burgett, Stanley G.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rao, Nagaraja R.
;    APPLICANT:  Richardson, Mark A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    TITLE OF INVENTION:  PLATENOLIDE SYNTHASE GENE
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PAUL R. CANTRELL  1138
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  Macintosh 7.0
;      SOFTWARE:  Microsoft Word 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,198
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CANTRELL, PAUL R.
;      REGISTRATION NUMBER:  36,470
;      REFERENCE/DOCKET NUMBER:  P9113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3885
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1580 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-804-198-5


  Query Match             46.8%;  Score 36;  DB 2;  Length 1580;
  Best Local Similarity   63.6%;  Pred. No. 3.1e+02;
  Matches    7;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 LAGTDFANQRP 14
           ||| |||   |
Db     752 LAGVDFAGHSP 762


RESULT 14
US-07-841-651-2
; Sequence 2, Application US/07841651
; Patent No. 5410031
;  GENERAL INFORMATION:
;    APPLICANT:  Pajor, Ana M
;    APPLICANT:  Wright, Ernest M
;    TITLE OF INVENTION:  Cloning and Functional Expression of a
;    TITLE OF INVENTION:  Mammalian Na+/Nucleoside Cotransporter: A Member of the
;    TITLE OF INVENTION:  SGLT Family
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheldon & Mak
;      STREET:  225 South Lake Avenue, Ninth Floor
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/841,651
;      FILING DATE:  19920224
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandel, SaraLynn
;      REGISTRATION NUMBER:  31,853
;      REFERENCE/DOCKET NUMBER:  8772
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (818) 796-4000
;      TELEFAX:  (818) 795-6321
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  672 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-841-651-2


  Query Match             45.5%;  Score 35;  DB 1;  Length 672;
  Best Local Similarity   70.0%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 IGLAGTDFAN 11
           :|||||  ||
Db      82 VGLAGTGAAN 91


RESULT 15
US-07-841-651-3
; Sequence 3, Application US/07841651
; Patent No. 5410031
;  GENERAL INFORMATION:
;    APPLICANT:  Pajor, Ana M
;    APPLICANT:  Wright, Ernest M
;    TITLE OF INVENTION:  Cloning and Functional Expression of a
;    TITLE OF INVENTION:  Mammalian Na+/Nucleoside Cotransporter: A Member of the
```

```
;    TITLE OF INVENTION:  SGLT Family
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheldon & Mak
;      STREET:  225 South Lake Avenue, Ninth Floor
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/841,651
;      FILING DATE:  19920224
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandel, SaraLynn
;      REGISTRATION NUMBER:  31,853
;      REFERENCE/DOCKET NUMBER:  8772
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (818) 796-4000
;      TELEFAX:  (818) 795-6321
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  672 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Oryctolagus cuniculus
US-07-841-651-3


  Query Match             45.5%;  Score 35;  DB 1;  Length 672;
  Best Local Similarity   70.0%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 IGLAGTDFAN 11
           :|||||  ||
Db      82 VGLAGTGAAN 91


Search completed: March 26, 2002, 13:41:29
Job time: 303 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:37:19 ; Search time 42.75 Seconds
(without alignments)
24.946 Million cell updates/sec

Title: US-09-709-201-100
Perfect score: 77
Sequence: 1 CIGLAGTDFANQRP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	68	88.3	392	2	A40371	major outer membra
2	43	55.8	995	2	T50267	probable family 31
3	41	53.2	1238	2	A64596	hypothetical prote
4	41	53.2	3194	2	D71917	toxin-like outer m
5	40	51.9	98	1	S32485	cytochrome c552 pr
6	39	50.6	289	2	S04407	phytoene synthase
7	39	50.6	324	2	C70785	probable carbohydr
8	39	50.6	377	3	JC7535	chitinase (EC 3.2.
9	39	50.6	502	2	A56683	receptor protein k
10	39	50.6	502	2	A53444	activin receptor-l
11	39	50.6	502	2	JC2491	serine/threonine k
12	39	50.6	660	2	JW0067	chitinase (EC 3.2.
13	39	50.6	743	2	G83726	assimilatory nitra
14	39	50.6	1785	2	T22595	hypothetical prote
15	38	49.4	356	2	F70636	probable ephB prot
16	38	49.4	365	2	T06615	hypothetical prote
17	37	48.1	141	2	JL0103	hypothetical 15.5K
18	37	48.1	276	2	S73410	hypothetical prote
19	37	48.1	547	2	S49814	transferrin-bindin
20	37	48.1	719	2	T33170	hypothetical prote
21	37	48.1	902	2	T49878	respiratory burst
22	37	48.1	926	2	E83375	probable glycosyl
23	37	48.1	1024	2	T46016	hypothetical prote
24	36.5	47.4	358	1	A48952	triacylglycerol li
25	36	46.8	155	2	D72696	riboflavin synthas
26	36	46.8	280	2	T24454	hypothetical prote
27	36	46.8	303	2	B70875	hypothetical prote
28	36	46.8	328	2	S74645	biliverdin reducta
29	36	46.8	354	2	A96596	hypothetical prote
30	36	46.8	371	2	T51695	cell division prot
31	36	46.8	416	2	S76310	hypothetical prote
32	36	46.8	464	2	B64970	hypothetical prote
33	36	46.8	464	2	C85830	hypothetical prote
34	36	46.8	467	2	S15297	hypothetical prote
35	36	46.8	513	2	T28933	hypothetical prote
36	36	46.8	616	2	S64624	alpha-glucoside tr
37	36	46.8	632	2	D71941	ATP-dependent zinc
38	36	46.8	632	2	E64653	cell division prot
39	36	46.8	638	2	T47267	cell cycle protein
40	36	46.8	645	2	G81315	membrane bound zin
41	36	46.8	725	2	T27148	hypothetical prote
42	36	46.8	1050	2	T31853	hypothetical prote
43	36	46.8	1060	2	T31763	hypothetical prote
44	35	45.5	101	2	C31982	Ca2+-transporting
45	35	45.5	104	2	T44890	hypothetical prote

ALIGNMENTS

RESULT 1
A40371
major outer membrane protein precursor - Chlamydophila psittaci (strain Fpn/pring)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 31-Mar-2000
C;Accession: I40859; A40371; S16137
R;Storey, C.; Lusher, M.; Yates, P.; Richmond, S.
J. Gen. Microbiol. 139, 2621-2626, 1993
A;Title: Evidence for Chlamydia pneumoniae of non-human origin.
A;Reference number: I40739; MUID:94103736
A;Accession: I40859
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A;Molecule type: DNA
A;Residues: 1-392 <RES>
A;Cross-references: EMBL:X61096; NID:g40564; PIDN:CAA43409.1; PID:g40565
A;Experimental source: strain Fpn
C;Genetics:
A;Gene: MOMP
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-392/Product: major outer membrane protein #status predicted <MAT>

Query Match 88.3%; Score 68; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 IGLAGTDFANQRP 14
           |||||||||||||
Db     160 IGLAGTDFANQRP 172
```

RESULT 2
T50267
probable family 31 glucosidase [imported] - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C;Accession: T50267
R;Hunt, C.; Aves, S.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1999
A;Reference number: Z25031
A;Accession: T50267
A;Status: preliminary; translated from GB/EMEL/DDBJ
A;Molecule type: DNA
A;Residues: 1-995 <HUN>
A;Cross-references: EMBL:AL133522; PIDN:CAB63549.1; GSPDB:GN00066; SPDB:SPAC922.02c
A;Experimental source: strain 972h(-); cosmid c922
C;Genetics:
A;Gene: SPAC1039.11c; SPDB:SPAC922.02c
A;Map position: 1
C;Superfamily: Schwanniomyces glucan 1,4-alpha-glucosidase GAM1; sucrase/isomaltase h

```
  Query Match             55.8%;  Score 43;  DB 2;  Length 995;
  Best Local Similarity   57.1%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDFANQRP 14
          |||  |||  :| |
Db    507 CIGSCGTDKLDQNP 520


RESULT   3
A64596
hypothetical protein HP0609 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C;Accession: A64596
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467
A;Accession: A64596
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1238 <TOM>
A;Cross-references: GB:AE000575; GB:AE000511; NID:g2313730; PIDN:AAD07677.1; PID:g231373


  Query Match             53.2%;  Score 41;  DB 2;  Length 1238;
  Best Local Similarity   53.8%;  Pred. No. 39;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          : |  ||| || |
Db    207 VNLTNTDFGNQTP 219


RESULT   4
D71917
toxin-like outer membrane protein jhp0556 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C;Accession: D71917
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557
A;Accession: D71917
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3194 <ARN>
A;Cross-references: GB:AE001488; GB:AE001439; NID:g4155100; PIDN:AAD06134.1; PID:g415510
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0556


  Query Match             53.2%;  Score 41;  DB 2;  Length 3194;
  Best Local Similarity   53.8%;  Pred. No. 1e+02;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          : |  ||| || |
Db    208 VNLTNTDFGNQTP 220


RESULT   5
S32485
```

```
cytochrome c552 precursor - Hydrogenobacter thermophilus
C;Species: Hydrogenobacter thermophilus
C;Date: 06-Jan-1995 #sequence_revision 27-Feb-1997 #text_change 03-Mar-2000
C;Accession: S32485; A32226
R;Sanbongi, Y.; Yang, J.H.; Igarashi, Y.; Kodama, T.
Eur. J. Biochem. 198, 7-12, 1991
A;Title: Cloning, nucleotide sequence and expression of the cytochrome c-552 gene fro
A;Reference number: S32485; MUID:91249816
A;Accession: S32485
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <SAN>
A;Cross-references: EMBL:X57735; NID:g43674; PIDN:CAA40902.1; PID:g43675
A;Experimental source: strain TK-6
R;Sanbongi, Y.; Ishii, M.; Igarashi, Y.; Kodama, T.
J. Bacteriol. 171, 65-69, 1989
A;Title: Amino acid sequence of cytochrome c-552 from a thermophilic hydrogen-oxidizi
A;Reference number: A32226; MUID:89123087
A;Accession: A32226
A;Molecule type: protein
A;Residues: 19-98 <SA2>
A;Experimental source: strain TK-6
C;Function:
A;Description: primary electron acceptor for molecular hydrogen activated by hydrogen
A;Pathway: hydrogen oxidation
C;Superfamily: cytochrome c6; cytochrome c6 homology
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein
F;1-18/Domain: signal sequence #status predicted <SIG>
F;17-94/Domain: cytochrome c6 homology <CYC>
F;19-98/Product: cytochrome c552 #status experimental <MAT>
F;28,31/Binding site: heme (Cys) (covalent) #status predicted
F;32,77/Binding site: heme iron (His, Met) (axial ligands) #status predicted


  Query Match             51.9%;  Score 40;  DB 1;  Length 98;
  Best Local Similarity   58.3%;  Pred. No. 4.7;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQR 13
          :|||| |||::
Db     10 VGLAGITFANEQ 21


RESULT   6
S04407
phytoene synthase - Rhodobacter capsulatus
C;Species: Rhodobacter capsulatus
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 08-Oct-1999
C;Accession: S04407
R;Armstrong, G.A.; Alberti, M.; Leach, F.; Hearst, J.E.
Mol. Gen. Genet. 216, 254-268, 1989
A;Title: Nucleotide sequence, organization, and nature of the protein products of the
A;Reference number: S04401; MUID:89313663
A;Accession: S04407
A;Molecule type: DNA
A;Residues: 1-289 <ARM>
A;Cross-references: EMBL:X52291; NID:g45996; PIDN:CAA36538.1; PID:g46003
C;Genetics:
A;Gene: crtE
C;Superfamily: geranyltranstransferase
C;Keywords: carotenoid biosynthesis


  Query Match             50.6%;  Score 39;  DB 2;  Length 289;
  Best Local Similarity   70.0%;  Pred. No. 21;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 AGTDFANQRP 14
          || | ||:||
Db    224 AGQDIANERP 233


RESULT   7
```

C70785
probable carbohydrate kinase - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: C70785
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: C70785
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-324 <COL>
A;Cross-references: GB:Z70283; GB:AL123456; NID:g3261561; PIDN:CAA94245.1; PID:g1237055
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: cbhK
C;Superfamily: probable ribokinase

Query Match 50.6%; Score 39; DB 2; Length 324;
Best Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IGLAGTDFANQR 13
:| || |||: |
Db 67 VGAAGADFADYR 78

RESULT 8
JC7535
chitinase (EC 3.2.1.14) 35 - Streptomyces thermoviolaceus
C;Species: Streptomyces thermoviolaceus
C;Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C;Accession: JC7535
R;Tsujibo, H.; Okamoto, T.; Hatano, N.; Miyamoto, K.; Watanabe, T.; Mitsutomi, M.; Inamo
Biosci. Biotechnol. Biochem. 64, 2445-2453, 2000
A;Title: Family 19 chitinases from Streptomyces thermoviolaceus OPC-520: Molecular cloni
A;Reference number: JC7535; MUID:21036907
A;Accession: JC7535
A;Molecule type: DNA
A;Residues: 1-377 <TSU>
A;Cross-references: DDBJ:AB016842
A;Experimental source: strain OPC-520
C;Comment: This enzyme, a member of the family 19 chitinases, is involved in chitin degr
lysaccharide-binding domain that is important in the efficient hydrolysis of insoluble c
C;Genetics:
A;Gene: chi35
C;Keywords: hydrolase; glycosidase

Query Match 50.6%; Score 39; DB 3; Length 377;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CIGLAGTDFANQRP 14
|: :|| | || |
Db 47 CLDVAGADSANGTP 60

RESULT 9
A56683
receptor protein kinase RPK-1 precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 24-Sep-1999
C;Accession: A56683
R;Sumitomo, S.; Saito, T.; Nohno, T.
DNA Seq. 3, 297-302, 1993
A;Title: A new receptor protein kinase from chick embryo related to type II receptor for
A;Reference number: A56683; MUID:94003400
A;Accession: A56683
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <SUM>
A;Cross-references: GB:D13432; NID:g222862; PIDN:BAA02694.1; PID:d1003199; PID:g22286
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C;Keywords: ATP; phosphotransferase; transmembrane protein
F;202-498/Domain: protein kinase homology <KIN>
F;210-218/Region: protein kinase ATP-binding motif

Query Match 50.6%; Score 39; DB 2; Length 502;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CIGLAGTDF 9
|:|| |:||
Db 71 CLGLEGSDF 79

RESULT 10
A53444
activin receptor-like kinase 6 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 24-Sep-1999
C;Accession: A53444; S40159
R;ten Dijke, P.; Yamashita, H.; Ichijo, H.; Franzen, P.; Laiho, M.; Miyazono, K.; Hel
Science 264, 101-104, 1994
A;Title: Characterization of type I receptors for transforming growth factor-beta and
A;Reference number: A53444; MUID:94188705
A;Accession: A53444
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <TEN>
A;Cross-references: GB:Z23143; NID:g437870; PIDN:CAA80674.1; PID:g437871
R;Miyazono, K.; Moren, A.; Grimsby, S.; Ichijo, H.; Heldin, C.; ten Dijke, P.
submitted to the EMBL Data Library, June 1993
A;Description: ALK-3 and ALK-6: the closely related members in the serine/threonine k
A;Reference number: S40158
A;Accession: S40159
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <MIY>
A;Cross-references: EMBL:Z23143; NID:g437870; PIDN:CAA80674.1; PID:g437871
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C;Keywords: ATP; serine/threonine-specific protein kinase; transmembrane protein
F;202-498/Domain: protein kinase homology <KIN>
F;210-218/Region: protein kinase ATP-binding motif

Query Match 50.6%; Score 39; DB 2; Length 502;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CIGLAGTDF 9
|:|| |:||
Db 71 CLGLEGSDF 79

RESULT 11
JC2491
serine/threonine kinase receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Apr-1995 #sequence_revision 26-May-1995 #text_change 10-Sep-1997
C;Accession: JC2491
R;Yamaji, N.; Celeste, A.J.; Thies, R.S.; Song, J.J.; Bernier, S.M.; Goltzman, D.; Ly
Biochem. Biophys. Res. Commun. 205, 1944-1951, 1994
A;Title: A mammalian serine/threonine kinase receptor specifically binds BMP-2 and BM
A;Reference number: JC2491; MUID:95110346
A;Accession: JC2491
A;Molecule type: mRNA
A;Residues: 1-502 <YAM>
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

```
C;Keywords: ATP; glycoprotein; transmembrane protein
F;127-148/Domain: transmembrane #status predicted <TMM>
F;202-498/Domain: protein kinase homology <KIN>
F;210-218/Region: protein kinase ATP-binding motif
F;284,343,388/Binding site: carbohydrate (Asn) (covalent) #status predicted


  Query Match           50.6%;  Score 39;  DB 2;  Length 502;
  Best Local Similarity  66.7%;  Pred. No. 36;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDF 9
          |:|| |:||
Db     71 CLGLEGSDF 79


RESULT  12
JW0067
chitinase (EC 3.2.1.14) A - Emericella nidulans
N;Alternate names: chiA
C;Species: Emericella nidulans, Aspergillus nidulans
C;Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 17-Mar-1999
C;Accession: JW0067
R;Takaya, N.; Yamazaki, D.; Horiuchi, H.; Ohta, A.; Takagi, M.
Biosci. Biotechnol. Biochem. 62, 60-65, 1998
A;Title: Cloning and characterization of a chitinase-encoding gene (chiA) from Aspergill
A;Reference number: JW0067; MUID:98162139
A;Accession: JW0067
A;Molecule type: mRNA
A;Residues: 1-660 <TAK>
A;Cross-references: DDBJ:D87895; NID:g2821948; PID:d1025495; PID:g2828335
C;Comment: This enzyme hydrolyzes chitin at belta-1,4 bonds between N-acetyglucosamine r
C;Genetics:
A;Gene: chiA
C;Keywords: glycosidase; hydrolase


  Query Match           50.6%;  Score 39;  DB 2;  Length 660;
  Best Local Similarity  70.0%;  Pred. No. 48;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 GLAGTDFANQ 12
          || |:|| ||
Db     71 GLPGSDFGNQ 80


RESULT  13
G83726
assimilatory nitrate reductase (catalytic subunit) nasC [imported] - Bacillus halodurans
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000
C;Accession: G83726
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20263314
A;Accession: G83726
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-743 <STO>
A;Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BAB04334.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: nasC
C;Superfamily: formate dehydrogenase


  Query Match           50.6%;  Score 39;  DB 2;  Length 743;
  Best Local Similarity  57.1%;  Pred. No. 54;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDFANQRP 14
          || |||| |  :|
Db    184 CIVLAGTNLAECQP 197


RESULT  14
T22595
hypothetical protein F53H4.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T22595
R;Dobson, R.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19587
A;Accession: T22595
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1785 <WIL>
A;Cross-references: EMBL:Z81089; PIDN:CAB03135.1; GSPDB:GN00028; CESP:F53H4.1
A;Experimental source: clone F53H4
C;Genetics:
A;Gene: CESP:F53H4.1
A;Map position: X
A;Introns: 42/2; 196/2; 245/2; 454/3; 562/2; 658/2; 730/3; 790/2; 844/3; 953/1; 1007/


  Query Match           50.6%;  Score 39;  DB 2;  Length 1785;
  Best Local Similarity  50.0%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDFANQRP 14
          |: | | | |  ||
Db    565 CVSLTGADSAAARP 578


RESULT  15
F70636
probable ephB protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: F70636
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A;Reference number: A70500; MUID:98295987
A;Accession: F70636
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-356 <COL>
A;Cross-references: GB:Z84498; GB:AL123456; NID:g3261701; PIDN:CAB06523.1; PID:g18062
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: ephB
C;Superfamily: tropinesterase


  Query Match           49.4%;  Score 38;  DB 2;  Length 356;
  Best Local Similarity  53.8%;  Pred. No. 39;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          ||| |: |  :||
Db    137 IGLPGSPFGERRP 149


Search completed: March 26, 2002, 13:37:21
Job time: 55 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:40:44 ; Search time 24.63 Seconds
(without alignments)
20.841 Million cell updates/sec

Title: US-09-709-201-100
Perfect score: 77
Sequence: 1 CIGLAGTDFANQRP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	68	88.3	392	1	OM1P_CHLPS	Q00087 chlamydia p
2	40	51.9	98	1	C552_HYDTH	P15452 hydrogenoba
3	39	50.6	289	1	CRTE_RHOCA	P17060 rhodobacter
4	39	50.6	324	1	YM02_MYCTU	Q10391 mycobacteri
5	39	50.6	502	1	BMRB_CHICK	Q05438 gallus gall
6	39	50.6	502	1	BMRB_HUMAN	O00238 homo sapien
7	39	50.6	502	1	BMRB_MOUSE	P36898 mus musculu
8	37	48.1	141	1	MMOB_METCA	P18797 methylococc
9	37	48.1	276	1	Y056_MYCPN	P75046 mycoplasma
10	36.5	47.4	358	1	LIP_PSEGL	Q05489 pseudomonas
11	36	46.8	155	1	RISC_AERPE	Q9ydc5 aeropyrum p
12	36	46.8	300	1	YGAF_SHIFL	P37775 shigella fl
13	36	46.8	426	1	GLYA_MYCLE	Q9x794 mycobacteri
14	36	46.8	464	1	WCAM_ECOLI	P71244 escherichia
15	36	46.8	467	1	WCAM_SALTY	P26389 salmonella
16	36	46.8	616	1	AGT1_YEAST	P53048 saccharomyc
17	36	46.8	632	1	FTSH_HELPJ	Q9zm66 helicobacte
18	36	46.8	632	1	FTSH_HELPY	P71408 helicobacte
19	36	46.8	638	1	FTSH_HELFE	O32617 helicobacte
20	35	45.5	118	1	GLBN_NOSSN	P52335 nostoc sp.
21	35	45.5	241	1	PSB1_YEAST	P23724 saccharomyc
22	35	45.5	273	1	ALKB_SCHPO	O60066 schizosacch
23	35	45.5	314	1	IUNH_CRIFA	Q27546 crithidia f
24	35	45.5	350	1	ADH1_CANAL	P43067 candida alb
25	35	45.5	390	1	IADA_ECOLI	P39377 escherichia
26	35	45.5	440	1	YHR1_YEAST	P38820 saccharomyc
27	35	45.5	496	1	SMVA_SALTY	P37594 salmonella
28	35	45.5	601	1	MAON_SOLTU	P37225 solanum tub
29	35	45.5	628	1	YZ68_PSEAE	P28812 pseudomonas
30	35	45.5	672	1	SL52_RABIT	P26430 oryctolagus
31	35	45.5	718	1	CTPC_MYCTU	P96875 mycobacteri
32	35	45.5	725	1	CTPC_MYCLE	Q9ccl1 mycobacteri
33	35	45.5	1112	1	ATB2_OREMO	P58165 oreochromis
34	35	45.5	1198	1	ATB2_MOUSE	Q9r0k7 mus musculu
35	35	45.5	1203	1	ATB4_RAT	Q64542 rattus norv
36	35	45.5	1220	1	ATB1_PIG	P23220 sus scrofa
37	35	45.5	1220	1	ATB3_HUMAN	Q16720 homo sapien
38	35	45.5	1241	1	ATB4_HUMAN	P23634 homo sapien
39	35	45.5	1243	1	ATB2_HUMAN	Q01814 homo sapien
40	35	45.5	1243	1	ATB2_RAT	P11506 rattus norv
41	35	45.5	1249	1	ATB1_RABIT	Q00804 oryctolagus
42	35	45.5	1258	1	ATB1_HUMAN	P20020 homo sapien
43	35	45.5	1258	1	ATB1_RAT	P11505 rattus norv
44	35	45.5	1258	1	ATB3_RAT	Q64568 rattus norv
45	35	45.5	1643	1	OMPB_RICPR	Q53020 r outer mem

ALIGNMENTS

```
RESULT   1
OM1P_CHLPS
ID   OM1P_CHLPS     STANDARD;      PRT;   392 AA.
AC   Q00087;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FPN/PRING;
RX   MEDLINE=94103736; PubMed=8277245;
RA   Storey C., Lusher M., Yates P., Richmond S.;
RT   "Evidence for Chlamydia pneumoniae of non-human origin.";
RL   J. Gen. Microbiol. 139:2621-2626(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61096; CAA43409.1; -.
DR   PIR; A40371; A40371.
DR   PIR; S16137; S16137.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    392       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   392 AA;  42069 MW;  88B3C5D90BBA26DB CRC64;


  Query Match             88.3%;  Score 68;  DB 1;  Length 392;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 IGLAGTDFANQRP 14
           |||||||||||||
Db     160 IGLAGTDFANQRP 172
```

```
RESULT   2
C552_HYDTH
ID   C552_HYDTH     STANDARD;      PRT;    98 AA.
AC   P15452;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CYTOCHROME C-552 PRECURSOR (C552).
OS   Hydrogenobacter thermophilus.
OC   Bacteria; Aquificales; Aquificaceae; Hydrogenobacter.
OX   NCBI_TaxID=940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TK-6;
RX   MEDLINE=91249816; PubMed=1645652;
RA   Sanbongi Y., Yang J.H., Igarashi Y., Kodama T.;
RT   "Cloning, nucleotide sequence and expression of the cytochrome c-552
RT   gene from Hydrogenobacter thermophilus.";
RL   Eur. J. Biochem. 198:7-12(1991).
RN   [2]
RP   SEQUENCE OF 19-98.
RC   STRAIN=TK-6;
RX   MEDLINE=89123087; PubMed=2536668;
RA   Sanbongi Y., Ishii M., Igarashi Y., Kodama T.;
RT   "Amino acid sequence of cytochrome c-552 from a thermophilic
RT   hydrogen-oxidizing bacterium, Hydrogenobacter thermophilus.";
RL   J. Bacteriol. 171:65-69(1989).
RN   [3]
RP   THERMOSTABILITY.
RX   MEDLINE=90122832; PubMed=2558725;
RA   Sanbongi Y., Igarashi Y., Kodama T.;
RT   "Thermostability of cytochrome c-552 from the thermophilic hydrogen-
RT   oxidizing bacterium Hydrogenobacter thermophilus.";
RL   Biochemistry 28:9574-9578(1989).
RN   [4]
RP   STRUCTURE BY NMR.
RC   STRAIN=TK-6;
RX   MEDLINE=98322065; PubMed=9657676;
RA   Hasegawa J., Yoshida T., Yamazaki T., Sambongi Y., Yu Y., Igarashi Y.,
RA   Kodama T., Yamazaki K., Kyogoku Y., Kobayashi Y.;
RT   "Solution structure of thermostable cytochrome c-552 from
RT   Hydrogenobacter thermophilus determined by 1H-NMR spectroscopy.";
RL   Biochemistry 37:9641-9649(1998).
CC   -!- FUNCTION: REACTS WITH HYDROGENASE.
CC   -!- PTM: BINDS ONE HEME GROUP PER MOLECULE.
CC   -!- SIMILARITY: 56% WITH P.AERUGINOSA C551.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X57735; CAA40902.1; -.
DR   PIR; A32226; A32226.
DR   PIR; S32485; S32485.
DR   PDB; 1AYG; 13-JAN-99.
DR   InterPro; IPR000345; CytC_heme_bind.
DR   InterPro; IPR003088; Cyt_CI.
DR   InterPro; IPR002324; Cyt_CID.
DR   Pfam; PF00034; cytochrome_c; 1.
DR   PRINTS; PR00606; CYTCHROMECID.
DR   PROSITE; PS00190; CYTOCHROME_C; 1.
KW   Electron transport; Heme; Signal; 3D-structure.
FT   SIGNAL        1     18
FT   CHAIN        19     98       CYTOCHROME C-552.
FT   BINDING      28     28       HEME (COVALENT).
FT   BINDING      31     31       HEME (COVALENT).
FT   METAL        32     32       IRON (HEME AXIAL LIGAND).
FT   METAL        77     77       IRON (HEME AXIAL LIGAND).
SQ   SEQUENCE   98 AA;  10431 MW;  F49713D829DDE927 CRC64;

  Query Match             51.9%;  Score 40;  DB 1;  Length 98;
  Best Local Similarity   58.3%;  Pred. No. 2;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 IGLAGTDFANQR 13
           :|||| |||::
Db      10 VGLAGITFANEQ 21


RESULT   3
CRTE_RHOCA
ID   CRTE_RHOCA     STANDARD;      PRT;   289 AA.
AC   P17060;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (EC 2.5.1.29) (GGPP
DE   SYNTHETASE) (FARNESYLTRANSTRANSFERASE).
GN   CRTE.
OS   Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Rhodobacter.
OX   NCBI_TaxID=1061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SB1003, AND BEC404;
RX   MEDLINE=89313663; PubMed=2747617;
RA   Armstrong G.A., Alberti M., Leach F., Hearst J.E.;
RT   "Nucleotide sequence, organization, and nature of the protein
RT   products of the carotenoid biosynthesis gene cluster of Rhodobacter
RT   capsulatus.";
RL   Mol. Gen. Genet. 216:254-268(1989).
CC   -!- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL
CC       DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.
CC   -!- PATHWAY: CAROTENOID AND CHLOROPHYLL BIOSYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52291; CAA36538.1; -.
DR   EMBL; Z11165; CAA77545.1; -.
DR   PIR; S04407; S04407.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Photosynthesis; Chlorophyll biosynthesis; Carotenoid biosynthesis;
KW   Isoprene biosynthesis; Transferase.
SQ   SEQUENCE   289 AA;  30043 MW;  CF483A26ECA9C859 CRC64;

  Query Match             50.6%;  Score 39;  DB 1;  Length 289;
  Best Local Similarity   70.0%;  Pred. No. 9.9;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 AGTDFANQRP 14
           || | ||:||
Db     224 AGQDIANERP 233


RESULT   4
YM02_MYCTU
ID   YM02_MYCTU     STANDARD;      PRT;   324 AA.
AC   Q10391;
DT   01-OCT-1996 (Rel. 34, Created)
```

```
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL SUGAR KINASE RV2202C.
GN   RV2202C OR MT2258.OR MTCY190.13C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE PFKB FAMILY OF CARBOHYDRATE KINASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z70283; CAA94245.1; -.
DR   EMBL; AE007071; AAK46544.1; -.
DR   TIGR; MT2258; -.
DR   TubercuList; Rv2202c; -.
DR   InterPro; IPR002173; PfkB.
DR   Pfam; PF00294; pfkB; 1.
DR   PROSITE; PS00583; PFKB_KINASES_1; 1.
DR   PROSITE; PS00584; PFKB_KINASES_2; FALSE_NEG.
KW   Hypothetical protein; Transferase; Kinase; Complete proteome.
SQ   SEQUENCE   324 AA;  34472 MW;  0C072206A3210A1D CRC64;

  Query Match           50.6%;  Score 39;  DB 1;  Length 324;
  Best Local Similarity  58.3%;  Pred. No. 11;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 IGLAGTDFANQR 13
           :| || |||: |
Db      67 VGAAGADFADYR 78


RESULT   5
BMRB_CHICK
ID   BMRB_CHICK     STANDARD;      PRT;   502 AA.
AC   Q05438;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   BONE MORPHOGENETIC PROTEIN RECEPTOR TYPE IB PRECURSOR (EC 2.7.1.37)
DE   (SERINE/THREONINE-PROTEIN KINASE RECEPTOR R6) (SKR6) (ACTIVIN
DE   RECEPTOR-LIKE KINASE 6) (ALK-6) (RPK-1).
GN   BMPR1B.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94003400; PubMed=8400359;
RA   Yamazaki Y., Saito T., Nohno T.;
RT   "A new receptor protein kinase from chick embryo related to type II
RT   receptor for TGF-beta.";
RL   DNA Seq. 3:297-302(1993).
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
CC   -!- SUBUNIT: HETERODIMERIZE WITH A TYPE-II RECEPTOR (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       TGFB RECEPTOR SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D13432; BAA02694.1; -.
DR   HSSP; P00523; 2PTK.
DR   InterPro; IPR000472; Activin_rec.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR003605; GS.
DR   InterPro; IPR002290; Ser_thr_kin_actsite.
DR   Pfam; PF01064; Activin_recp; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   SMART; SM00467; GS; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Transmembrane; Glycoprotein; Signal.
FT   SIGNAL       1     13       POTENTIAL.
FT   CHAIN       14    502       BONE MORPHOGENETIC PROTEIN RECEPTOR TYPE
FT                               IB.
FT   DOMAIN      14    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    127   148       POTENTIAL.
FT   DOMAIN      149   502       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      204   494       PROTEIN KINASE.
FT   NP_BIND     210   218       ATP (BY SIMILARITY).
FT   BINDING     231   231       ATP (BY SIMILARITY).
FT   ACT_SITE    332   332       BY SIMILARITY.
FT   CARBOHYD     44    44       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   502 AA;  56766 MW;  D5D93CCEBF2A0680 CRC64;

  Query Match           50.6%;  Score 39;  DB 1;  Length 502;
  Best Local Similarity  66.7%;  Pred. No. 18;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 CLGLEGSDF 79


RESULT   6
BMRB_HUMAN
ID   BMRB_HUMAN     STANDARD;      PRT;   502 AA.
AC   O00238; P78366;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
```

```
DE   BONE MORPHOGENETIC PROTEIN RECEPTOR TYPE IB PRECURSOR (EC 2.7.1.37).
GN   BMPR1B.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=97322244; PubMed=9178898;
RA   Ide H., Katoh M., Sasaki H., Yoshida T., Aoki K., Nawa Y., Osada Y.,
RA   Sugimura T., Terada M.;
RT   "Cloning of human bone morphogenetic protein type IB receptor (BMPR-
RT   IB) and its expression in prostate cancer in comparison with other
RT   BMPRs.";
RL   Oncogene 14:1377-1382(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Astrom A.-K., Jin D.F., Imamura T., Roijer E., Rosenzweig B.,
RA   Miyazono K., ten Dijke P., Stenman G.;
RT   "Chromosomal localization of three human genes encoding bone
RT   morphogenetic protein receptors.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR BMPS/OP-1.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
CC   -!- SUBUNIT: HETERODIMERIZE WITH A TYPE-II RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       TGFB RECEPTOR SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D89675; BAA19765.1; -.
DR   EMBL; U89326; AAC28131.1; -.
DR   MIM; 603248; -.
DR   InterPro; IPR000333; ActivinII_rcptor.
DR   InterPro; IPR000472; Activin_rec.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR003605; GS.
DR   InterPro; IPR002290; Ser_thr_kin_actsite.
DR   InterPro; IPR001245; Tyr_kin.
DR   Pfam; PF01064; Activin_recp; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00653; ACTIVIN2R.
DR   SMART; SM00467; GS; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Transmembrane; Signal.
FT   SIGNAL       1     13       POTENTIAL.
FT   CHAIN       14    502       BONE MORPHOGENETIC PROTEIN RECEPTOR TYPE
FT                               IB.
FT   DOMAIN      14    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   127    148       POTENTIAL.
FT   DOMAIN     149    502       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     204    494       PROTEIN KINASE.
FT   NP_BIND    210    218       ATP (BY SIMILARITY).
FT   BINDING    231    231       ATP (BY SIMILARITY).
FT   ACT_SITE   332    332       BY SIMILARITY.
SQ   SEQUENCE   502 AA;  56930 MW;  B283D9BF45535C79 CRC64;

  Query Match           50.6%;  Score 39;  DB 1;  Length 502;
  Best Local Similarity 66.7%;  Pred. No. 18;

    Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 CLGLEGSDF 79


RESULT   7
BMRB_MOUSE
ID   BMRB_MOUSE     STANDARD;      PRT;   502 AA.
AC   P36898;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   BONE MORPHOGENETIC PROTEIN RECEPTOR TYPE IB PRECURSOR (EC 2.7.1.37)
DE   (SERINE/THREONINE-PROTEIN KINASE RECEPTOR R6) (SKR6) (ACTIVIN
DE   RECEPTOR-LIKE KINASE 6) (ALK-6).
GN   BMPR1B OR ACVRLK6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94188705; PubMed=8140412;
RA   ten Dijke P., Yamashita H., Ichijo H., Franzen P., Laiho M.,
RA   Miyazono K., Heldin C.-H.;
RT   "Characterization of type I receptors for transforming growth factor-
RT   beta and activin.";
RL   Science 264:101-104(1994).
CC   -!- FUNCTION: RECEPTOR FOR BMPS/OP-1.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
CC   -!- SUBUNIT: HETERODIMERIZE WITH A TYPE-II RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       TGFB RECEPTOR SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z23143; CAA80674.1; -.
DR   PIR; A53444; A53444.
DR   PIR; S40159; S40159.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:107191; Bmprlb.
DR   InterPro; IPR000472; Activin_rec.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR003605; GS.
DR   InterPro; IPR002290; Ser_thr_kin_actsite.
DR   Pfam; PF01064; Activin_recp; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   SMART; SM00467; GS; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Transmembrane; Signal.
FT   SIGNAL       1     13       POTENTIAL.
FT   CHAIN       14    502       BONE MORPHOGENETIC PROTEIN RECEPTOR TYPE
FT                               IB.
FT   DOMAIN      14    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   127    148       POTENTIAL.
FT   DOMAIN     149    502       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     204    494       PROTEIN KINASE.
FT   NP_BIND    210    218       ATP (BY SIMILARITY).
FT   BINDING    231    231       ATP (BY SIMILARITY).
FT   ACT_SITE   332    332       BY SIMILARITY.
SQ   SEQUENCE   502 AA;  56944 MW;  AB29681F3FF5A361 CRC64;
```

```
  Query Match             50.6%;  Score 39;  DB 1;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 18;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDF 9
          |:|| |:||
Db     71 CLGLEGSDF 79


RESULT   8
MMOB_METCA
ID   MMOB_METCA     STANDARD;      PRT;   141 AA.
AC   P18797;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-AUG-1991 (Rel. 19, Last annotation update)
DE   METHANE MONOOXYGENASE REGULATORY PROTEIN B.
GN   MMOB.
OS   Methylococcus capsulatus.
OC   Bacteria; Proteobacteria; gamma subdivision; Methylococcaceae;
OC   Methylococcus.
OX   NCBI_TaxID=414;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BATH;
RX   MEDLINE=89373399; PubMed=2505721;
RA   Stainthorpe A.C., Murrell J.C., Salmond G.P.C., Dalton H., Lees V.;
RT   "Molecular analysis of methane monooxygenase from Methylococcus
RT   capsulatus (Bath).";
RL   Arch. Microbiol. 152:154-159(1989).
RN   [2]
RP   IDENTIFICATION OF PROTEIN, AND SEQUENCE OF 3-19.
RA   Pilkington S.J., Salmond G.P.C., Murrell J.C., Dalton H.;
RT   "Identification of the gene encoding the regulatory protein B of
RT   soluble methane monooxygenase.";
RL   FEMS Microbiol. Lett. 72:345-348(1990).
CC   -!- FUNCTION: THE B PROTEIN ACTS AS A REGULATOR OF ELECTRON FLOW
CC       THROUGH THE SOLUBLE MMO COMPLEX, SWITCHING THE ENZYME FROM AN
CC       OXIDASE TO A HYDROXYLASE IN THE PRESENCE OF THE SUBSTRATE.
CC   -!- SUBUNIT: M.CAPSULATUS HAS TWO FORMS OF METHANE MONOOXYGENASE,
CC       A SOLUBLE AND A MEMBRANE-BOUND TYPE. THE SOLUBLE TYPE CONSISTS
CC       OF THREE COMPONENTS (A, B AND C).
DR   PIR; JL0103; JL0103.
DR   InterPro; IPR003454; MmoB_DmpM.
DR   Pfam; PF02406; MmoB_DmpM; 1.
KW   Oxidoreductase; Monooxygenase.
SQ   SEQUENCE   141 AA;  16018 MW;  45A42A61D6C58406 CRC64;


  Query Match             48.1%;  Score 37;  DB 1;  Length 141;
  Best Local Similarity   63.6%;  Pred. No. 11;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQ 12
          :|| | |||:|
Db     13 MGLKGKDFADQ 23


RESULT   9
Y056_MYCPN
ID   Y056_MYCPN     STANDARD;      PRT;   276 AA.
AC   P75046;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL PROTEIN MG056 HOMOLOG (D09_ORF276).
GN   MPN071 OR MP084.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
CC   -!- SIMILARITY: BELONGS TO THE UPF0011 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000010; AAB95732.1; -.
DR   InterPro; IPR000878; Corrin_porph_mthyltrnf.
DR   InterPro; IPR000578; UPF0011.
DR   Pfam; PF00590; TP_methylase; 1.
DR   ProDom; PD007098; UPF0011; 1.
DR   PROSITE; PS01296; UPF0011; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   276 AA;  31062 MW;  2A071E7A17FCD0CD CRC64;


  Query Match             48.1%;  Score 37;  DB 1;  Length 276;
  Best Local Similarity   50.0%;  Pred. No. 22;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQR 13
          :|| | ||:|::
Db     43 LGLLGIDFSNKQ 54


RESULT  10
LIP_PSEGL
ID   LIP_PSEGL      STANDARD;      PRT;   353 AA.
AC   Q05489;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   LIPASE PRECURSOR (EC 3.1.1.3) (TRIACYLGLYCEROL LIPASE).
GN   LIPA.
OS   Pseudomonas glumae, and Chromobacterium viscosum.
OC   Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC   Burkholderia.
OX   NCBI_TaxID=337, 42739;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 40-61, AND MUTAGENESIS.
RC   SPECIES=P.glumae; STRAIN=PG1 / CBS 322.89;
RX   MEDLINE=93119130; PubMed=1476423;
RA   Frenken L.G.J., Egmond M.R., Batenburg A.M., Bos J.W., Visser C.,
RA   Verrips C.T.;
RT   "Cloning of the Pseudomonas glumae lipase gene and determination of
RT   the active site residues.";
RL   Appl. Environ. Microbiol. 58:3787-3791(1992).
RN   [2]
RP   SEQUENCE OF 40-54, AND CHARACTERIZATION.
RC   SPECIES=C.viscosum;
RX   MEDLINE=95306500; PubMed=7786905;
RA   Taipa M.A., Liebeton K., Costa J.V., Cabral J.M.S., Jaeger K.-E.;
RT   "Lipase from Chromobacterium viscosum: biochemical characterization
RT   indicating homology to the lipase from Pseudomonas glumae.";
RL   Biochim. Biophys. Acta 1256:396-402(1995).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RC   SPECIES=P.glumae;
RX   MEDLINE=94009622; PubMed=8405390;
RA   Noble M.E.M., Cleasby A., Johnson L.N., Egmond M.R., Frenken L.G.J.;
```

```
RT   "The crystal structure of triacylglycerol lipase from Pseudomonas
RT   glumae reveals a partially redundant catalytic aspartate.";
RL   FEBS Lett. 331:123-128(1993).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RC   SPECIES=C.viscosum; STRAIN=ATCC 6918;
RX   MEDLINE=96275656; PubMed=8683577;
RA   Lang D., Hofmann B., Haalck L., Hecht H.-J., Spener F., Schmid R.D.,
RA   Schomburg D.;
RT   "Crystal structure of a bacterial lipase from Chromobacterium
RT   viscosum ATCC 6918 refined at 1.6-A resolution.";
RL   J. Mol. Biol. 259:704-717(1996).
CC   -!- FUNCTION: HYDROLYSIS OF TRIGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +
CC       A FATTY ACID ANION.
CC   -!- COFACTOR: REQUIRES CALCIUM.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: STRONG TO OTHER PSEUDOMONAS LIPASES.
CC   -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC       HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X70354; CAA49812.1; -.
DR   EMBL; A16323; CAA01279.1; -.
DR   EMBL; A32021; CAA02073.1; -.
DR   PIR; A48952; A48952.
DR   PIR; S37291; S37291.
DR   PDB; 1TAH; 31-MAY-94.
DR   InterPro; IPR000073; Abhydrolase.
DR   InterPro; IPR000379; Est_lip_thioest_actsite.
DR   InterPro; IPR000734; Lipase.
DR   Pfam; PF00561; abhydrolase; 1.
DR   PROSITE; PS00120; LIPASE_SER; 1.
KW   Hydrolase; Lipid degradation; Signal; Calcium; 3D-structure.
FT   SIGNAL       1     39
FT   CHAIN       40    358       LIPASE.
FT   ACT_SITE    126   126       CHARGE RELAY SYSTEM.
FT   ACT_SITE    302   302       CHARGE RELAY SYSTEM.
FT   ACT_SITE    324   324       CHARGE RELAY SYSTEM.
FT   DISULFID    229   308
FT   MUTAGEN      54    54       H->A: NO LOSS OF ACTIVITY.
FT   MUTAGEN     126   126       S->A: COMPLETE LOSS OF ACTIVITY.
FT   MUTAGEN     160   160       D->E: NO LOSS OF ACTIVITY.
FT   MUTAGEN     160   160       D->A: NO LOSS OF ACTIVITY.
FT   MUTAGEN     280   280       D->E: NO LOSS OF ACTIVITY.
FT   MUTAGEN     280   280       D->A: COMPLETE LOSS OF ACTIVITY.
FT   MUTAGEN     302   302       D->E: NO LOSS OF ACTIVITY.
FT   MUTAGEN     302   302       D->A: 75% LOSS OF ACTIVITY.
FT   MUTAGEN     324   324       H->A: COMPLETE LOSS OF ACTIVITY.
FT   CONFLICT     40    40       A -> W (IN REF. 2).
SQ   SEQUENCE   358 AA;  36928 MW;  FE7B5D7A22EC6B4B CRC64;


  Query Match           47.4%;  Score 36.5;  DB 1;  Length 358;
  Best Local Similarity  90.0%;  Pred. No. 36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       3 GLAGTD-FAN 11
           |||||| |||
Db      55 GLAGTDKFAN 64


RESULT  11
RISC_AERPE
ID   RISC_AERPE     STANDARD;      PRT;   155 AA.
AC   Q9YDC5;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   RIBOFLAVIN SYNTHASE (EC 2.5.1.9).
GN   RIBC OR APE0988.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC   Aeropyrum.
OX   NCBI_TaxID=56636;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1;
RX   MEDLINE=99310339; PubMed=10382966;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA   Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
RT   crenarchaeon, Aeropyrum pernix K1.";
RL   DNA Res. 6:83-101(1999).
CC   -!- CATALYTIC ACTIVITY: 2 6,7-DIMETHYL-8-(1-D-RIBITYL)LUMAZINE =
CC       RIBOFLAVIN + 4-(1-D-RIBITYLAMINO)-5-AMINO-2,6-DIHYDROXYPYRIMIDINE.
CC   -!- COFACTOR: FLAVOPROTEIN (BY SIMILARITY).
CC   -!- PATHWAY: FINAL STEP OF RIBOFLAVIN SYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AP000060; BAA79972.1; -.
KW   Riboflavin biosynthesis; Transferase; Flavoprotein; Complete proteome.
SQ   SEQUENCE   155 AA;  16576 MW;  644B7A1BD313A4F3 CRC64;


  Query Match           46.8%;  Score 36;  DB 1;  Length 155;
  Best Local Similarity  60.0%;  Pred. No. 18;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDFA 10
           |:|:| | ||
Db       4 CVGVADTTFA 13


RESULT  12
YGAF_SHIFL
ID   YGAF_SHIFL     STANDARD;      PRT;   300 AA.
AC   P37775;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   HYPOTHETICAL PROTEIN IN GALF 5'REGION (ORF1X3) (FRAGMENT).
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SEROTYPE 2A;
RX   MEDLINE=94131953; PubMed=7507920;
RA   Morona R., Mavris M., Fallarino A., Manning P.A.;
RT   "Characterization of the rfc region of Shigella flexneri.";
RL   J. Bacteriol. 176:733-747(1994).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X71970; CAA50765.1; -.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   300 AA;  33367 MW;  4735B2EF93E21325 CRC64;


  Query Match             46.8%;  Score 36;  DB 1;  Length 300;
  Best Local Similarity   53.8%;  Pred. No. 36;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 IGLAGTDFANQRP 14
           |||||: : |  |
Db      73 IGLAGSTYDNNYP 85


RESULT  13
GLYA_MYCLE
ID   GLYA_MYCLE     STANDARD;      PRT;   426 AA.
AC   Q9X794;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DE   (SHMT).
GN   GLYA OR ML1953 OR MLCB1222.22.
OS   Mycobacterium leprae.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA   Murphy L., Oliver K., Quail M.A., Rajandream M.-A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
CC   -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC   -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC       H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC       HORMONES AND OTHER COMPONENTS.
CC   -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AL049491; CAB39828.1; -.
DR   EMBL; AL583923; CAC30908.1; -.
DR   Leproma; ML1953; -.
DR   InterPro; IPR001085; SHMT.
DR   Pfam; PF00464; SHMT; 1.
DR   PROSITE; PS00096; SHMT; FALSE_NEG.
KW   Transferase; Pyridoxal phosphate; One-carbon metabolism;
KW   Complete proteome.
FT   BINDING     227     227       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   426 AA;  45224 MW;  27783E2328AF2C98 CRC64;


  Query Match             46.8%;  Score 36;  DB 1;  Length 426;
  Best Local Similarity   63.6%;  Pred. No. 53;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 LAGTDFANQRP 14
           | | |||| :|
Db      82 LFGADFANVQP 92


RESULT  14
WCAM_ECOLI
ID   WCAM_ECOLI     STANDARD;      PRT;   464 AA.
AC   P71244; P76378;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   COLANIC ACID BIOSYNTHESIS PROTEIN WCAM.
GN   WCAM OR B2043.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=96326333; PubMed=8759852;
RA   Stevenson G., Andrianopoulos K., Hobbs M., Reeves P.R.;
RT   "Organization of the Escherichia coli K-12 gene cluster responsible
RT   for production of the extracellular polysaccharide colanic acid.";
RL   J. Bacteriol. 178:4885-4893(1996).
RN   [2]
RP   REVISIONS.
RC   STRAIN=K12;
RA   Reeves P.R.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA   Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
RA   Saito N., Sampei G., Seki Y., Sivasundaram S., Tagami H.,
RA   Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
CC   -!- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF THE SLIME POLYSACCHARIDE
CC       COLANIC ACID.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U38473; AAC77852.1; -.
DR   EMBL; AE000295; AAC75104.1; -.
DR   EMBL; D90842; BAA15897.1; -.
DR   EcoGene; EG12651; wcaM.
KW   Lipopolysaccharide biosynthesis; Complete proteome.
SQ   SEQUENCE   464 AA;  51315 MW;  72A7655DC07368BE CRC64;


  Query Match             46.8%;  Score 36;  DB 1;  Length 464;
  Best Local Similarity   53.8%;  Pred. No. 58;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          |||||: : |  |
Db    228 IGLAGSTYDNNYP 240


RESULT  15
WCAM_SALTY
ID   WCAM_SALTY     STANDARD;      PRT;   467 AA.
AC   P26389;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   COLANIC ACID BIOSYNTHESIS PROTEIN WCAM.
GN   WCAM.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=91260454; PubMed=1710759;
RA   Jiang X.-M., Neal B., Santiago F., Lee S.J., Romana L.K., Reeves P.R.;
RT   "Structure and sequence of the rfb (O antigen) gene cluster of
RT   Salmonella serovar typhimurium (strain LT2).";
RL   Mol. Microbiol. 5:695-713(1991).
CC   -!- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF THE SLIME POLYSACCHARIDE
CC       COLANIC ACID.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X56793; CAA40113.1; -.
DR   PIR; S15297; S15297.
DR   StyGene; SG10448; wcaM.
KW   Lipopolysaccharide biosynthesis.
SQ   SEQUENCE   467 AA;  50958 MW;  9DCCFD551218E6E8 CRC64;


  Query Match             46.8%;  Score 36;  DB 1;  Length 467;
  Best Local Similarity   53.8%;  Pred. No. 58;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          |||||: : |  |
Db    231 IGLAGSTYDNSYP 243


Search completed: March 26, 2002, 13:40:45
Job time: 259 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:40:14 ; Search time 79.01 Seconds
(without alignments)
25.918 Million cell updates/sec

Title: US-09-709-201-100
Perfect score: 77
Sequence: 1 CIGLAGTDFANQRP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	68	88.3	392	2	Q99QB0	Q99qb0 chlamydophi
2	50	64.9	356	2	O52924	O52924 chlamydia p
3	50	64.9	390	2	Q9AIJ5	Q9aij5 chlamydia p
4	50	64.9	392	2	Q9AIJ4	Q9aij4 chlamydia p
5	43	55.8	188	3	Q9HFA9	Q9hfa9 trichosporo
6	43	55.8	995	3	Q9URX4	Q9urx4 schizosacch
7	41	53.2	1238	2	O25330	O25330 helicobacte
8	41	53.2	2546	2	Q9AI30	Q9ai30 burkholderi
9	41	53.2	3194	2	Q9ZLM3	Q9zlm3 helicobacte
10	39	50.6	377	2	Q9RHU5	Q9rhu5 streptomyce
11	39	50.6	382	2	Q9A111	Q9a111 streptococc
12	39	50.6	502	6	Q9BDI4	Q9bdi4 ovis aries
13	39	50.6	502	11	Q9QVT7	Q9qvt7 rattus sp.
14	39	50.6	502	13	Q9PUF4	Q9puf4 coturnix co
15	39	50.6	743	2	Q9KF71	Q9kf71 bacillus ha
16	39	50.6	961	3	Q92223	Q92223 emericella
17	39	50.6	1785	5	Q93781	Q93781 caenorhabdi
18	38	49.4	356	2	P95276	P95276 mycobacteri
19	38	49.4	365	10	Q9SZ73	Q9sz73 arabidopsis
20	38	49.4	388	3	Q9HDZ5	Q9hdz5 schizosacch
21	38	49.4	543	13	Q9W645	Q9w645 gallus gall
22	38	49.4	564	2	Q9F4V4	Q9f4v4 photorhabdu
23	38	49.4	614	10	Q9LV47	Q9lv47 arabidopsis
24	38	49.4	640	5	O61604	O61604 drosophila
25	37	48.1	141	2	Q9LCT7	Q9lct7 methylococc
26	37	48.1	144	2	Q9ACS1	Q9acs1 streptomyce
27	37	48.1	270	13	Q9I9C0	Q9i9c0 trachemys s
28	37	48.1	293	2	Q9ZJ02	Q9zj02 streptococc
29	37	48.1	479	2	Q9KZE9	Q9kze9 streptomyce
30	37	48.1	500	5	Q9VPK7	Q9vpk7 drosophila
31	37	48.1	539	4	Q9NY99	Q9ny99 homo sapien
32	37	48.1	547	2	Q44167	Q44167 actinobacil
33	37	48.1	553	2	Q99SS5	Q99ss5 staphylococ
34	37	48.1	619	5	P91944	P91944 drosophila
35	37	48.1	719	5	O61862	O61862 caenorhabdi
36	37	48.1	719	5	Q9W4B8	Q9w4b8 drosophila
37	37	48.1	902	10	O81209	O81209 arabidopsis
38	37	48.1	902	10	Q9LY21	Q9ly21 arabidopsis
39	37	48.1	926	2	Q9I1V3	Q9i1v3 pseudomonas
40	37	48.1	1024	10	Q9M2Q4	Q9m2q4 arabidopsis
41	36.5	47.4	687	5	Q9VMJ5	Q9vmj5 drosophila
42	36	46.8	187	4	Q9H4Z9	Q9h4z9 homo sapien
43	36	46.8	187	12	Q66033	Q66033 cercopithec
44	36	46.8	205	2	Q9FA16	Q9fa16 acetobacter
45	36	46.8	263	5	Q9GS30	Q9gs30 drosophila

ALIGNMENTS

RESULT 1
Q99QB0
ID Q99QB0 PRELIMINARY; PRT; 392 AA.
AC Q99QB0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN OMPA.
OS Chlamydophila felis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP BAKER, ATCC VR120, AND FP CELLO;
RX MEDLINE=21078680; PubMed=11211261;
RA Bush R.M., Everett K.D.;
RT "Molecular evolution of the Chlamydiaceae.";
RL Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR EMBL; AF269257; AAK00238.1; -.
DR EMBL; AF269258; AAK00239.1; -.
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 392 MAJOR OUTER MEMBRANE PROTEIN.
SQ SEQUENCE 392 AA; 42051 MW; 88B3C09C1FEE26DB CRC64;

Query Match 88.3%; Score 68; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 0.00059;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IGLAGTDFANQRP 14
||||||||||||||
Db 160 IGLAGTDFANQRP 172

RESULT 2
O52924
ID O52924 PRELIMINARY; PRT; 356 AA.
AC O52924;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER MEMBRANE PROTEIN 1 (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=92-1293, AVIAN SEROVAR D;
RA   Vanrompay D., Cox E., Goddeeris B.M., Volckaert G.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y16562; CAA76286.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane.
FT   NON_TER     356    356
SQ   SEQUENCE   356 AA;  38396 MW;  D51DE06FB46E6F13 CRC64;


  Query Match             64.9%;  Score 50;  DB 2;  Length 356;
  Best Local Similarity   76.9%;  Pred. No. 0.8;
  Matches   10;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          ||| |||| || |
Db    160 IGLKGTDFNNQLP 172


RESULT   3
Q9AIJ5
ID   Q9AIJ5      PRELIMINARY;      PRT;   390 AA.
AC   Q9AIJ5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW JERSEY 1, NJ1;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269266; AAK00247.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     20       POTENTIAL.
FT   CHAIN        21    390       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   390 AA;  42042 MW;  B62858403DBFA4E6 CRC64;


  Query Match             64.9%;  Score 50;  DB 2;  Length 390;
  Best Local Similarity   76.9%;  Pred. No. 0.88;
  Matches   10;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          ||| |||| || |
Db    158 IGLKGTDFNNQLP 170


RESULT   4
Q9AIJ4
ID   Q9AIJ4      PRELIMINARY;      PRT;   392 AA.
AC   Q9AIJ4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TEXAS TURKEY 3, TT3;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269267; AAK00248.1; -.
KW   Signal.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23    392       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   392 AA;  42293 MW;  FC31FC051955246C CRC64;


  Query Match             64.9%;  Score 50;  DB 2;  Length 392;
  Best Local Similarity   76.9%;  Pred. No. 0.88;
  Matches   10;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          ||| |||| || |
Db    160 IGLKGTDFNNQLP 172


RESULT   5
Q9HFA9
ID   Q9HFA9      PRELIMINARY;      PRT;   188 AA.
AC   Q9HFA9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   20 KDA PROTEIN HAVING G-X-X-X-Q-X-W-MOTIF PRECURSOR.
OS   Trichosporon mucoides.
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Tremellomycetidae;
OC   Trichosporonales; Trichosporon.
OX   NCBI_TaxID=82522;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIMM 1573;
RA   Usui Y., Matsunaga Y.;
RT   "20 kDa protein having G-X-X-X-Q-X-W motif.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB038375; BAB20766.1; -.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF00652; Ricin_B_lectin; 3.
DR   SMART; SM00458; RICIN; 1.
KW   Signal.
FT   SIGNAL       1     16       POTENTIAL.
SQ   SEQUENCE   188 AA;  19751 MW;  A391173C35A547B1 CRC64;


  Query Match             55.8%;  Score 43;  DB 3;  Length 188;
  Best Local Similarity   50.0%;  Pred. No. 7;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDFANQRP 14
          |: :|| :|||  |
Db     57 CVDVAGANFANGTP 70


RESULT   6
Q9URX4
ID   Q9URX4      PRELIMINARY;      PRT;   995 AA.
AC   Q9URX4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE FAMILY 31 GLUCOSIDASE.
GN   SPAC922.02C.
OS   Schizosaccharomyces pombe (Fission yeast).
```

```
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL133522; CAB63549.1; -.
DR   InterPro; IPR000322; Glyco_hydro_31.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF01055; Glyco_hydro_31; 1.
DR   PROSITE; PS00129; GLYCOSYL_HYDROL_F31_1; 1.
DR   PROSITE; PS00707; GLYCOSYL_HYDROL_F31_2; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
SQ   SEQUENCE   995 AA;  112713 MW;  1EC1D292DC30DBA8 CRC64;


  Query Match             55.8%;  Score 43;  DB 3;  Length 995;
  Best Local Similarity   57.1%;  Pred. No. 40;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDFANQRP 14
          |||  |||  :| |
Db    507 CIGSCGTDKLDQNP 520


RESULT   7
O25330
ID   O25330      PRELIMINARY;      PRT;  1238 AA.
AC   O25330;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   HYPOTHETICAL 135.1 KDA PROTEIN.
GN   HP0609.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
DR   EMBL; AE000575; AAD07677.1; -.
DR   TIGR; HP0609; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   1238 AA;  135062 MW;  66DF754EB1BFB173 CRC64;


  Query Match             53.2%;  Score 41;  DB 2;  Length 1238;
  Best Local Similarity   53.8%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          : |  ||| || |
Db    207 VNLTNTDFGNQTP 219


RESULT   8
Q9AI30
ID   Q9AI30      PRELIMINARY;      PRT;  2546 AA.
AC   Q9AI30;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE TYPE I POLYKETIDE SYNTHASE WCBR.
GN   WCBR.
OS   Burkholderia mallei (Pseudomonas mallei).
OC   Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC   Burkholderia.
OX   NCBI_TaxID=13373;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC23344;
RA   DeShazer D., Waag D.M., Fritz D.L., Woods D.E.;
RT   "Identification of a Burkholderia mallei polysaccharide gene cluster
RT   by subtractive hybridization and demonstration that the encoded
RT   capsule is an essential virulence determinant.";
RL   Microb. Pathog. 0:0-0(2001).
DR   EMBL; AF285636; AAK26474.1; -.
SQ   SEQUENCE   2546 AA;  267597 MW;  55DFD9BC44A5F9BA CRC64;


  Query Match             53.2%;  Score 41;  DB 2;  Length 2546;
  Best Local Similarity   54.5%;  Pred. No. 2.4e+02;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQ 12
          :|:||||: |:
Db    123 VGVAGTDYGNR 133


RESULT   9
Q9ZLM3
ID   Q9ZLM3      PRELIMINARY;      PRT;  3194 AA.
AC   Q9ZLM3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE VACUOLATING CYTOTOXIN (VACA) PARALOG.
GN   JHP0556.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
DR   EMBL; AE001488; AAD06134.1; -.
DR   InterPro; IPR001589; Actinin_act_bind.
DR   PROSITE; PS00019; ACTININ_1; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   3194 AA;  348350 MW;  26D60C492DBECF0E CRC64;


  Query Match             53.2%;  Score 41;  DB 2;  Length 3194;
  Best Local Similarity   53.8%;  Pred. No. 3e+02;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 IGLAGTDFANQRP 14
          : |  ||| || |
Db    208 VNLTNTDFGNQTP 220
```

```
RESULT  10
Q9RHU5
ID   Q9RHU5      PRELIMINARY;      PRT;   377 AA.
AC   Q9RHU5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CHI35 PROTEIN.
GN   CHI35.
OS   Streptomyces thermoviolaceus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1952;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OPC-520;
RA   Tsujibo H.;
RT   "Chitinase.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB016842; BAA88833.1; -.
DR   HSSP; P23951; 2BAA.
DR   InterPro; IPR000726; Glyco_hydro_19.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF00182; Glyco_hydro_19; 1.
DR   Pfam; PF00652; Ricin_B_lectin; 1.
DR   ProDom; PD000574; Glyco_hydro_19; 1.
DR   SMART; SM00458; RICIN; 1.
SQ   SEQUENCE   377 AA;  39763 MW;  14267B344738562B CRC64;

  Query Match             50.6%;  Score 39;  DB 2;  Length 377;
  Best Local Similarity   50.0%;  Pred. No. 74;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDFANQRP 14
           |: :|| | ||  |
Db      47 CLDVAGADSANGTP 60


RESULT  11
Q9A111
ID   Q9A111      PRELIMINARY;      PRT;   382 AA.
AC   Q9A111;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE ACETYL-COA C-ACETYLTRANSFERASE (EC 2.3.1.).
GN   SPY0524.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR   EMBL; AE006510; AAK33520.1; -.
KW   Transferase; Acyltransferase; Complete proteome.
SQ   SEQUENCE   382 AA;  40746 MW;  67755D330D73C2D4 CRC64;

  Query Match             50.6%;  Score 39;  DB 2;  Length 382;
  Best Local Similarity   53.8%;  Pred. No. 75;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 IGLAGTDFANQRP 14
           ||| |  || ::|
Db      14 IGLVGKQFAKEQP 26


RESULT  12
Q9BDI4
ID   Q9BDI4      PRELIMINARY;      PRT;   502 AA.
AC   Q9BDI4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   BONE MORPHOGENETIC PROTEIN RECEPTOR TYPE IB.
GN   BMPR-IB.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Souza C.J.H., MacDougall C.N., Campbell B.K., McNeilly A.S.,
RA   Baird D.T.;
RT   "The Booroola (FecB) phenotype is associated with a mutation in the
RT   bone morphogenetic receptor type 1B (BMPR1B) gene.";
RL   J. Endocrinol. 0:0-0(2001).
DR   EMBL; AF357007; AAK30296.1; -.
KW   Receptor.
SQ   SEQUENCE   502 AA;  56907 MW;  6552124A0A24F35C CRC64;

  Query Match             50.6%;  Score 39;  DB 6;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 99;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CIGLAGTDF 9
           |:|| |:||
Db      71 CLGLEGSDF 79


RESULT  13
Q9QVT7
ID   Q9QVT7      PRELIMINARY;      PRT;   502 AA.
AC   Q9QVT7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CFK-43A=BONE MORPHOGENETIC PROTEIN BINDING SERINE/THREONINE KINASE
DE   RECEPTOR_.
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95110346; PubMed=7811286;
RA   Yamaji N., Celeste A.J., Thies R.S., Song J.J., Bernier S.M.,
RA   Goltzman D., Lyons K.M., Nove J., Rosen V., Wozney J.M.;
RT   "A mammalian serine/threonine kinase receptor specifically binds BMP-2
RT   and BMP-4.";
RL   Biochem. Biophys. Res. Commun. 205:1944-1951(1994).
CC   -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   InterPro; IPR000333; ActivinII_receptor.
DR   InterPro; IPR000472; Activin_rec.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR003605; GS.
DR   InterPro; IPR002290; Ser_thr_kin_actsite.
DR   InterPro; IPR001245; Tyr_kin.
DR   Pfam; PF01064; Activin_recp; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00653; ACTIVIN2R.
DR   PRINTS; PR00109; TYRKINASE.
DR   SMART; SM00467; GS; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
```

```
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   502 AA;  56870 MW;  E147D1B4477F7573 CRC64;


  Query Match            50.6%;  Score 39;  DB 11;  Length 502;
  Best Local Similarity  66.7%;  Pred. No. 99;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDF 9
          |:|| |:||
Db     71 CLGLEGSDF 79


RESULT  14
Q9PUF4
ID   Q9PUF4      PRELIMINARY;      PRT;   502 AA.
AC   Q9PUF4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   BONE MORPHOGENETIC PROTEIN RECEPTOR IB.
GN   BMPR-IB.
OS   Coturnix coturnix (common quail).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Coturnix.
OX   NCBI_TaxID=9091;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McPherson C.E., Varley J.E., Maxwell G.D.;
RT   "Expression and Regulation of the Type I BMP Receptors during Avian
RT   Neural Crest Development.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   EMBL; AF189778; AAF04583.1; -.
DR   HSSP; P36897; 1TBI.
DR   InterPro; IPR000333; ActivinII_receptor.
DR   InterPro; IPR000472; Activin_rec.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR003605; GS.
DR   InterPro; IPR002290; Ser_thr_kin_actsite.
DR   InterPro; IPR001245; Tyr_kin.
DR   Pfam; PF01064; Activin_recp; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00653; ACTIVIN2R.
DR   PRINTS; PR00109; TYRKINASE.
DR   SMART; SM00467; GS; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Receptor; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   502 AA;  56886 MW;  6819265085F28422 CRC64;


  Query Match            50.6%;  Score 39;  DB 13;  Length 502;
  Best Local Similarity  66.7%;  Pred. No. 99;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDF 9
          |:|| |:||
Db     71 CLGLEGSDF 79


RESULT  15
Q9KF71
ID   Q9KF71      PRELIMINARY;      PRT;   743 AA.
AC   Q9KF71;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ASSIMILATORY NITRATE REDUCTASE (CATALYTIC SUBUNIT).
GN   NASC OR BH0615.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001509; BAB04334.1; -.
DR   InterPro; IPR001467; Molybdopterin.
DR   Pfam; PF00384; molybdopterin; 1.
DR   Pfam; PF01568; Molydop_binding; 1.
DR   PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; UNKNOWN_1.
DR   PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   743 AA;  82692 MW;  C7221E52CB270E1A CRC64;


  Query Match            50.6%;  Score 39;  DB 2;  Length 743;
  Best Local Similarity  57.1%;  Pred. No. 1.5e+02;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CIGLAGTDFANQRP 14
          || |||| |  :|
Db    184 CIVLAGTNLAECQP 197


Search completed: March 26, 2002, 13:40:15
Job time: 229 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:38:48 ; Search time 81.51 Seconds
(without alignments)
11.814 Million cell updates/sec

Title: US-09-709-201-101
Perfect score: 73
Sequence: 1 CQINKFKSRKACG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73	100.0	13	20	AAW95328	Peptide fragment o
2	73	100.0	343	20	AAY56771	C. trachomatis ser
3	73	100.0	391	20	AAY35319	Chlamydia pneumoni
4	67	91.8	93	20	AAW95319	Chlamydial major o
5	65	89.0	13	20	AAW95330	Peptide fragment o
6	65	89.0	13	20	AAW95331	Peptide fragment o
7	59	80.8	13	20	AAW95329	Peptide fragment o
8	58	79.5	215	20	AAW98185	Chlamydia psittaci
9	58	79.5	222	20	AAW98183	Chlamydia psittaci
10	58	79.5	225	20	AAW98186	Chlamydia psittaci
11	58	79.5	228	20	AAW98184	Chlamydia psittaci
12	58	79.5	343	20	AAY56769	C. trachomatis ser
13	58	79.5	356	20	AAY56770	C. trachomatis ser
14	58	79.5	389	20	AAW98188	Chlamydia psittaci
15	58	79.5	402	20	AAW98189	Chlamydia psittaci
16	58	79.5	402	20	AAW98187	Chlamydia psittaci
17	58	79.5	525	21	AAB13645	C. pneumoniae sero
18	58	79.5	525	22	AAG83213	Protein encoded by
19	53	72.6	94	20	AAW95308	Chlamydial major o
20	53	72.6	94	20	AAW95310	Chlamydial major o
21	53	72.6	94	20	AAW95311	Chlamydial major o
22	53	72.6	94	20	AAW95312	Chlamydial major o
23	53	72.6	94	20	AAW95315	Chlamydial major o
24	53	72.6	94	20	AAW95316	Chlamydial major o
25	53	72.6	95	20	AAW95313	Chlamydial major o
26	53	72.6	95	20	AAW95314	Chlamydial major o
27	53	72.6	95	20	AAW95317	Chlamydial major o
28	53	72.6	277	21	AAY82393	C. trachomatis MOM
29	53	72.6	349	21	AAY82392	C. trachomatis MOM
30	53	72.6	372	19	AAW76365	C. trachomatis JM1
31	53	72.6	373	19	AAW76362	C. trachomatis JM1
32	53	72.6	374	19	AAW76364	C. trachomatis JM1
33	53	72.6	376	19	AAW76363	C. trachomatis JM1
34	53	72.6	376	19	AAW76366	C. trachomatis JM1
35	53	72.6	387	20	AAY56767	C. trachomatis ser
36	53	72.6	392	20	AAY56760	C. trachomatis ser
37	53	72.6	393	20	AAY56757	C. trachomatis ser
38	53	72.6	393	20	AAY56759	C. trachomatis ser
39	53	72.6	393	22	AAE06646	Chlamydia trachoma
40	53	72.6	394	7	AAP60004	Sequence of a majo
41	53	72.6	394	18	AAW15149	Chlamydia trachoma
42	53	72.6	394	19	AAW73141	Chlamydia trachoma
43	53	72.6	394	19	AAW57775	Chlamydia trachoma
44	53	72.6	394	20	AAY56758	C. trachomatis ser
45	53	72.6	394	21	AAY81268	Chlamydia trachoma

ALIGNMENTS

RESULT 1
AAW95328
ID AAW95328 standard; Protein; 13 AA.
XX
AC AAW95328;
XX
DT 15-MAR-1999 (first entry)
XX
DE Peptide fragment of C. pneumoniae CPN342-354.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia pneumoniae.
XX
PN WO9850074-A2.
XX
PD 12-NOV-1998.
XX
PF 06-MAY-1998; 98WO-US09237.
XX
PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
PR 18-FEB-1998; 98US-0025176.
XX

PA (UYVA-) UNIV VANDERBILT.
XX
PI Mitchell WM, Stratton CW;
XX
DR WPI; 1999-059653/05.
XX
PT Composition with two agents effective against different stages of
PT chlamydial life cycle - comprises agent targetted against cryptic
PT phase, against elementary body phase, against replicating phase,
PT probenicid and antiporphyric
XX
PS Claim 4; Fig 5; 138pp; English.
XX
CC The invention relates to the diagnosis and management of infections by
CC Chlamydia species. The invention provides a composition that comprises
CC at least two agents, where each of the agents is effective against a
CC different phase of the chlamydial life cycle. The agents are selected
CC from: (a) agents targetted against cryptic phase of chlamydial life
CC cycle; (b) agents targetted against elementary body phase of chlamydial
CC life cycle; (c) agents targetted against replicating phase of chlamydial
CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC is used to elicit a protective immune response to Chlamydia infection in
CC an animal or human and is applied until the animal or human tests
CC negative for Chlamydia infection. It is also used to treat biological
CC material infected with Chlamydia. Diagnostic kits for antibody assays
CC against recombinant major outer membrane protein (MOMP), and for DNA
CC amplification assays for chlamydial genes, are used to diagnose disease,
CC e.g. autoimmune disease, an inflammatory disease or a disease that
CC occurs in an immuno-compromised individual, associated with Chlamydia
CC infection. The kits are used to detect chlamydial elementary bodies in a
CC sample. They are also used to monitor and/or modify the course of therapy
CC in a patient. The treatment reduces the acellular load of infectious
CC Ebstein Barr virus. The method is also used to treat porphyria, by
CC reducing the number of elementary bodies and applying a drug, e.g.
CC cimetidine, and antioxidants, to reduce the adverse effects associated
CC with porphyria. Sequences AAW95328 to AAW95331 represent peptides
CC employed for the construction of peptide based ELISAs with species
CC specificity for variable domain 2 (VD2).
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 73; DB 20; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQINKFKSRKACG 13
|||||||||||||
Db 1 cqinkfksrkacg 13

RESULT 2
AAY56771
ID AAY56771 standard; Protein; 343 AA.
XX
AC AAY56771;
XX
DT 22-FEB-2000 (first entry)
XX
DE C. trachomatis serovar HuPn MOMP sequence.
XX
KW Major outer membrane protein; MOMP; Chlamydia; vaccine; immune response;
KW cellular response; immunogen; Th1-like CD4 response; mucosal immunity.
XX
OS Chlamydia trachomatis.
XX
PN WO9951745-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-CA00292.
XX
PR 07-APR-1998; 98US-0055765.
XX
PA (UYMA-) UNIV MANITOBA.
XX
PI Bruhnam RC;
XX
DR WPI; 1999-620205/53.
XX
PT Non-replicating vector encoding fragments of the outer membrane protein
PT of Chlamydia, useful in vaccines and as immunogen -
XX
PS Disclosure; Fig 10 A-F; 52pp; English.
XX
CC The invention provides a non-replicating vector that comprises, linked
CC to a promoter, a nucleotide sequence that encodes a region containing at
CC least one of the conserved domains 2, 3 and 5 of a major outer membrane
CC protein (MOMP) of a Chlamydia strain. The vector is used: (a) in
CC vaccines to generate a protective immune response (mainly cellular)
CC against MOMP, and (b) as immunogens to raise anti-MOMP antibodies, useful
CC in standard immunoassays. Immunization with the vector induces a broad
CC spectrum of immune responses, including Th1-like CD4 responses and
CC mucosal immunity, providing significant protection against subsequent
CC challenge. Sequences AAY56757-71 represent MOMP sequences from a variety
CC of serovars of C. trachomatis.
XX
SQ Sequence 343 AA;

Query Match 100.0%; Score 73; DB 20; Length 343;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQINKFKSRKACG 13
|||||||||||||
Db 296 cqinkfksrkacg 308

RESULT 3
AAY35319
ID AAY35319 standard; Protein; 391 AA.
XX
AC AAY35319;
XX
DT 13-SEP-1999 (first entry)
XX
DE Chlamydia pneumoniae transmembrane protein sequence.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-IB01890.
XX
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
PA (GEST) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae
XX
PS Page 1130-1131; Disclosure; 1912pp; English.
XX
CC AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.

CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotides sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 391 AA;

Query Match 100.0%; Score 73; DB 20; Length 391;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQINKFKSRKACG 13
||||||||||||
Db 344 cqinkfksrkacg 356

RESULT 4
AAW95319
ID AAW95319 standard; Protein; 93 AA.
XX
AC AAW95319;
XX
DT 15-MAR-1999 (first entry)
XX
DE Chlamydial major outer membrane protein (MOMP) PN fragment.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia sp.
XX
PN WO9850074-A2.
XX
PD 12-NOV-1998.
XX
PF 06-MAY-1998; 98WO-US09237.
XX
PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
PR 18-FEB-1998; 98US-0025176.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Mitchell WM, Stratton CW;
XX
DR WPI; 1999-059653/05.
XX
PT Composition with two agents effective against different stages of
PT chlamydial life cycle - comprises agent targetted against cryptic
PT phase, against elementary body phase, against replicating phase,
PT probenicid and antiporphyric
XX
PS Disclosure; Fig 1A; 138pp; English.
XX
CC The invention relates to the diagnosis and management of infections by
CC Chlamydia species. The invention provides a composition that comprises
CC at least two agents, where each of the agents is effective against a
CC different phase of the chlamydial life cycle. The agents are selected
CC from: (a) agents targetted against cryptic phase of chlamydial life
CC cycle; (b) agents targetted against elementary body phase of chlamydial
CC life cycle; (c) agents targetted against replicating phase of chlamydial
CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC is used to elicit a protective immune response to Chlamydia infection in
CC an animal or human and is applied until the animal or human tests
CC negative for Chlamydia infection. It is also used to treat biological
CC material infected with Chlamydia. Diagnostic kits for antibody assays
CC against recombinant major outer membrane protein (MOMP), and for DNA
CC amplification assays for chlamydial genes, are used to diagnose disease,
CC e.g. autoimmune disease, an inflammatory disease or a disease that
CC occurs in an immuno-compromised individual, associated with Chlamydia
CC infection. The kits are used to detect chlamydial elementary bodies in a
CC sample. They are also used to monitor and/or modify the course of therapy
CC in a patient. The treatment reduces the acellular load of infectious
CC Ebstein Barr virus. The method is also used to treat porphyria, by
CC reducing the number of elementary bodies and applying a drug, e.g.
CC cimetidine, and antioxidants, to reduce the adverse effects associated
CC with porphyria. Sequences AAW95272 to AAW95319 represent peptide
CC fragments of various Chlamydial MOMPs.
XX
SQ Sequence 93 AA;

Query Match 91.8%; Score 67; DB 20; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQINKFKSRKAC 12
||||||||||||
Db 46 cqinkfksrkac 57

RESULT 5
AAW95330
ID AAW95330 standard; Protein; 13 AA.
XX
AC AAW95330;
XX
DT 15-MAR-1999 (first entry)
XX
DE Peptide fragment of C. trachomatis CTL342-354.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia trachomatis.
XX
PN WO9850074-A2.
XX
PD 12-NOV-1998.
XX
PF 06-MAY-1998; 98WO-US09237.
XX
PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
PR 18-FEB-1998; 98US-0025176.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Mitchell WM, Stratton CW;
XX
DR WPI; 1999-059653/05.
XX

PT Composition with two agents effective against different stages of
PT chlamydial life cycle - comprises agent targetted against cryptic
PT phase, against elementary body phase, against replicating phase,
PT probenicid and antiporphyric
XX
PS Claim 4; Fig 5; 138pp; English.
XX
CC The invention relates to the diagnosis and management of infections by
CC Chlamydia species. The invention provides a composition that comprises
CC at least two agents, where each of the agents is effective against a
CC different phase of the chlamydial life cycle. The agents are selected
CC from: (a) agents targetted against cryptic phase of chlamydial life
CC cycle; (b) agents targetted against elementary body phase of chlamydial
CC life cycle; (c) agents targetted against replicating phase of chlamydial
CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC is used to elicit a protective immune response to Chlamydia infection in
CC an animal or human and is applied until the animal or human tests
CC negative for Chlamydia infection. It is also used to treat biological
CC material infected with Chlamydia. Diagnostic kits for antibody assays
CC against recombinant major outer membrane protein (MOMP), and for DNA
CC amplification assays for chlamydial genes, are used to diagnose disease,
CC e.g. autoimmune disease, an inflammatory disease or a disease that
CC occurs in an immuno-compromised individual, associated with Chlamydia
CC infection. The kits are used to detect chlamydial elementary bodies in a
CC sample. They are also used to monitor and/or modify the course of therapy
CC in a patient. The treatment reduces the acellular load of infectious
CC Ebstein Barr virus. The method is also used to treat porphyria, by
CC reducing the number of elementary bodies and applying a drug, e.g.
CC cimetidine, and antioxidants, to reduce the adverse effects associated
CC with porphyria. Sequences AAW95328 to AAW95331 represent peptides
CC employed for the construction of peptide based ELISAs with species
CC specificity for variable domain 2 (VD2).
XX
SQ Sequence 13 AA;

Query Match 89.0%; Score 65; DB 20; Length 13;
Best Local Similarity 84.6%; Pred. No. 0.00012;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CQINKFKSRKACG 13
||:|| |||||||
Db 1 cqlnkmksrkacg 13

RESULT 6
AAW95331
ID AAW95331 standard; Protein; 13 AA.
XX
AC AAW95331;
XX
DT 15-MAR-1999 (first entry)
XX
DE Peptide fragment of C. psitacci CPS342-354.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia psitacci.
XX
PN WO9850074-A2.
XX
PD 12-NOV-1998.
XX
PF 06-MAY-1998; 98WO-US09237.
XX
PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
PR 18-FEB-1998; 98US-0025176.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Mitchell WM, Stratton CW;
XX
DR WPI; 1999-059653/05.
XX
PT Composition with two agents effective against different stages of
PT chlamydial life cycle - comprises agent targetted against cryptic
PT phase, against elementary body phase, against replicating phase,
PT probenicid and antiporphyric
XX
PS Claim 4; Fig 5; 138pp; English.
XX
CC The invention relates to the diagnosis and management of infections by
CC Chlamydia species. The invention provides a composition that comprises
CC at least two agents, where each of the agents is effective against a
CC different phase of the chlamydial life cycle. The agents are selected
CC from: (a) agents targetted against cryptic phase of chlamydial life
CC cycle; (b) agents targetted against elementary body phase of chlamydial
CC life cycle; (c) agents targetted against replicating phase of chlamydial
CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC is used to elicit a protective immune response to Chlamydia infection in
CC an animal or human and is applied until the animal or human tests
CC negative for Chlamydia infection. It is also used to treat biological
CC material infected with Chlamydia. Diagnostic kits for antibody assays
CC against recombinant major outer membrane protein (MOMP), and for DNA
CC amplification assays for chlamydial genes, are used to diagnose disease,
CC e.g. autoimmune disease, an inflammatory disease or a disease that
CC occurs in an immuno-compromised individual, associated with Chlamydia
CC infection. The kits are used to detect chlamydial elementary bodies in a
CC sample. They are also used to monitor and/or modify the course of therapy
CC in a patient. The treatment reduces the acellular load of infectious
CC Ebstein Barr virus. The method is also used to treat porphyria, by
CC reducing the number of elementary bodies and applying a drug, e.g.
CC cimetidine, and antioxidants, to reduce the adverse effects associated
CC with porphyria. Sequences AAW95328 to AAW95331 represent peptides
CC employed for the construction of peptide based ELISAs with species
CC specificity for variable domain 2 (VD2).
XX
SQ Sequence 13 AA;

Query Match 89.0%; Score 65; DB 20; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.00012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CQINKFKSRKACG 13
||||||||| |||
Db 1 cqinkfksrfacg 13

RESULT 7
AAW95329
ID AAW95329 standard; Protein; 13 AA.
XX
AC AAW95329;
XX
DT 15-MAR-1999 (first entry)
XX
DE Peptide fragment of C. trachomatis CTP342-354.
XX
KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyric acid; imune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX
OS Chlamydia trachomatis.

```
XX
PN   WO9850074-A2.
XX
PD   12-NOV-1998.
XX
PF   06-MAY-1998;   98WO-US09237.
XX
PR   18-FEB-1998;   98US-0025521.
PR   06-MAY-1997;   97US-0045689.
PR   06-MAY-1997;   97US-0045739.
PR   06-MAY-1997;   97US-0045779.
PR   06-MAY-1997;   97US-0045780.
PR   06-MAY-1997;   97US-0045784.
PR   06-MAY-1997;   97US-0045787.
PR   14-AUG-1997;   97US-0911593.
PR   18-FEB-1998;   98US-0025174.
PR   18-FEB-1998;   98US-0025176.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Mitchell WM,  Stratton CW;
XX
DR   WPI; 1999-059653/05.
XX
PT   Composition with two agents effective against different stages of
PT   chlamydial life cycle - comprises agent targetted against cryptic
PT   phase, against elementary body phase, against replicating phase,
PT   probenicid and antiporphyric
XX
PS   Claim 4; Fig 5; 138pp; English.
XX
CC   The invention relates to the diagnosis and management of infections by
CC   Chlamydia species. The invention provides a composition that comprises
CC   at least two agents, where each of the agents is effective against a
CC   different phase of the chlamydial life cycle. The agents are selected
CC   from: (a) agents targetted against cryptic phase of chlamydial life
CC   cycle; (b) agents targetted against elementary body phase of chlamydial
CC   life cycle; (c) agents targetted against replicating phase of chlamydial
CC   life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC   is used to elicit a protective immune response to Chlamydia infection in
CC   an animal or human and is applied until the animal or human tests
CC   negative for Chlamydia infection. It is also used to treat biological
CC   material infected with Chlamydia. Diagnostic kits for antibody assays
CC   against recombinant major outer membrane protein (MOMP), and for DNA
CC   amplification assays for chlamydial genes, are used to diagnose disease,
CC   e.g. autoimmune disease, an inflammatory disease or a disease that
CC   occurs in an immuno-compromised individual, associated with Chlamydia
CC   infection. The kits are used to detect chlamydial elementary bodies in a
CC   sample. They are also used to monitor and/or modify the course of therapy
CC   in a patient. The treatment reduces the acellular load of infectious
CC   Ebstein Barr virus. The method is also used to treat porphyria, by
CC   reducing the number of elementary bodies and applying a drug, e.g.
CC   cimetidine, and antioxidants, to reduce the adverse effects associated
CC   with porphyria. Sequences AAW95328 to AAW95331 represent peptides
CC   employed for the construction of peptide based ELISAs with species
CC   specificity for variable domain 2 (VD2).
XX
SQ   Sequence   13 AA;


  Query Match             80.8%;  Score 59;  DB 20;  Length 13;
  Best Local Similarity   84.6%;  Pred. No. 0.0013;
  Matches   11;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CQINKFKSRKACG 13
           ||||| ||| |||
Db       1 cqinkmksrfacg 13


RESULT   8
AAW98185
ID   AAW98185 standard; Protein; 215 AA.
XX
AC   AAW98185;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Chlamydia psittaci MOMP (minus VD1 and VD2 region).
XX
KW   Major outer membrane protein; MOMP; psittacosis; infection;
KW   vaccine; genetic immunisation.
XX
OS   Chlamydia psittaci.
XX
PN   WO9910005-A1.
XX
PD   04-MAR-1999.
XX
PF   28-AUG-1998;   98WO-US17943.
XX
PR   28-AUG-1997;   97US-0057147.
XX
PA   (LOUU ) UNIV LOUISIANA & AGRIC & MECH COLLEGE.
XX
PI   Baghian A,  Chouljenko VN,  Kousoulas KG,  Tully TN;
XX
DR   WPI; 1999-254214/21.
XX
PT   A new vaccine for Chlamydia psittaci infections
XX
PS   Claim 28; Page 51; 72pp; English.
XX
CC   The present sequence is a major outer membrane protein (MOMP)
CC   polypeptide of the cockatiel isolate LSUWTCK of Chlamydia psittaci
CC   (the MOMP gene sequence of this isolate is identical to that of C.
CC   psittaci Avian Type C).  The MOMP polypeptide comprises regions VD3
CC   and VD4 of native MOMP (see also AAW98137), i.e. it lacks regions VD1
CC   and VD2 of MOMP.  A claimed method of preventing a C. psittaci
CC   infection in a subject comprises administering an immunising
CC   amount of an expression vector comprising a eukaryotic promoter
CC   functionally linked to a nucleic acid encoding a C. psittaci
CC   MOMP polypeptide lacking regions VD1 and VD2, preferably the
CC   present sequence or a polypeptide (see also AAW98186) from C.
CC   psittaci strain B577.
XX
SQ   Sequence   215 AA;


  Query Match             79.5%;  Score 58;  DB 20;  Length 215;
  Best Local Similarity   91.7%;  Pred. No. 0.027;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     169 qinkmksrkacg 180


RESULT   9
AAW98183
ID   AAW98183 standard; Protein; 222 AA.
XX
AC   AAW98183;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Chlamydia psittaci MOMP (minus VD1 and VD2 region).
XX
KW   Major outer membrane protein; MOMP; psittacosis; infection;
KW   vaccine; genetic immunisation.
XX
OS   Chlamydia psittaci.
XX
PN   WO9910005-A1.
XX
PD   04-MAR-1999.
XX
```

PF 28-AUG-1998; 98WO-US17943.
XX
PR 28-AUG-1997; 97US-0057147.
XX
PA (LOUU) UNIV LOUISIANA & AGRIC & MECH COLLEGE.
XX
PI Baghian A, Chouljenko VN, Kousoulas KG, Tully TN;
XX
DR WPI; 1999-254214/21.
DR N-PSDB; AAX25044.
XX
PT A new vaccine for Chlamydia psittaci infections
XX
PS Claim 6; Page 41-42; 72pp; English.
XX
CC The present sequence is a major outer membrane protein (MOMP)
CC polypeptide of Chlamydia psittaci strain LSUWTCK, a cockatiel
CC isolate (the MOMP gene sequence of this isolate is identical to
CC that of C. psittaci Avian Type C). The polypeptide comprises
CC regions VD3 and VD4 of native MOMP (see also AAW98187), i.e. it lacks
CC regions VD1 and VD2 of MOMP. A claimed vaccine composition includes
CC the MOMP polypeptide, optionally fused to a maltose binding protein.
CC Also claimed are an isolated nucleic acid (see AAX25044) encoding the
CC polypeptide, a vector, and a method of preventing C. psittaci
CC infection by administering the vaccine containing the MOMP
CC polypeptide. Vectors encoding MOMP polypeptides lacking regions
CC VD1 and VD2 are useful for genetic (naked nucleic acid) vaccination.
CC The vaccines are used to prevent C. psittaci infection, especially
CC in birds.
XX
SQ Sequence 222 AA;

Query Match 79.5%; Score 58; DB 20; Length 222;
Best Local Similarity 91.7%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QINKFKSRKACG 13
|||| |||||||
Db 176 qinkmksrkacg 187

RESULT 10
AAW98186
ID AAW98186 standard; Protein; 225 AA.
XX
AC AAW98186;
XX
DT 05-JUL-1999 (first entry)
XX
DE Chlamydia psittaci MOMP (minus VD1 and VD2 region).
XX
KW Major outer membrane protein; MOMP; psittacosis; infection;
KW vaccine; genetic immunisation.
XX
OS Chlamydia psittaci.
XX
PN WO9910005-A1.
XX
PD 04-MAR-1999.
XX
PF 28-AUG-1998; 98WO-US17943.
XX
PR 28-AUG-1997; 97US-0057147.
XX
PA (LOUU) UNIV LOUISIANA & AGRIC & MECH COLLEGE.
XX
PI Baghian A, Chouljenko VN, Kousoulas KG, Tully TN;
XX
DR WPI; 1999-254214/21.
XX
PT A new vaccine for Chlamydia psittaci infections
XX
PS Claim 28; Page 51-52; 72pp; English.
XX
CC The present sequence is a major outer membrane protein (MOMP)
CC polypeptide of Chlamydia psittaci strain B577. It comprises
CC regions VD3 and VD4 of native MOMP (see also AAW98187), but lacks
CC regions VD1 and VD2. A claimed method of preventing a C. psittaci
CC infection in a subject comprises administering an immunising
CC amount of an expression vector comprising a eukaryotic promoter
CC functionally linked to a nucleic acid encoding a C. psittaci
CC MOMP polypeptide lacking regions VD1 and VD2, preferably the
CC present sequence or a polypeptide (see also AAW98186) from C.
CC psittaci strain B577.
XX
SQ Sequence 225 AA;

Query Match 79.5%; Score 58; DB 20; Length 225;
Best Local Similarity 91.7%; Pred. No. 0.029;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QINKFKSRKACG 13
|||| |||||||
Db 179 qinkmksrkacg 190

RESULT 11
AAW98184
ID AAW98184 standard; Protein; 228 AA.
XX
AC AAW98184;
XX
DT 05-JUL-1999 (first entry)
XX
DE Chlamydia psittaci MOMP (minus VD1 and VD2 region).
XX
KW Major outer membrane protein; MOMP; psittacosis; infection;
KW vaccine; genetic immunisation.
XX
OS Chlamydia psittaci.
XX
PN WO9910005-A1.
XX
PD 04-MAR-1999.
XX
PF 28-AUG-1998; 98WO-US17943.
XX
PR 28-AUG-1997; 97US-0057147.
XX
PA (LOUU) UNIV LOUISIANA & AGRIC & MECH COLLEGE.
XX
PI Baghian A, Chouljenko VN, Kousoulas KG, Tully TN;
XX
DR WPI; 1999-254214/21.
DR N-PSDB; AAX25045.
XX
PT A new vaccine for Chlamydia psittaci infections
XX
PS Claim 6; Page 42-43; 72pp; English.
XX
CC The present sequence is a major outer membrane protein (MOMP)
CC polypeptide of Chlamydia psittaci strain B577. The polypeptide
CC comprises regions VD3 and VD4 of the native protein (see also
CC AAW98187), but lacks regions VD1 and VD2. A claimed vaccine
CC composition includes this MOMP polypeptide, optionally fused to a
CC maltose binding protein. Also claimed are an isolated nucleic
CC acid (see AAX25045) encoding the polypeptide, a vector, and a method
CC of preventing C. psittaci infection by administering the vaccine
CC containing the MOMP polypeptide. Vectors encoding MOMP polypeptides
CC lacking regions VD1 and VD2 are useful for genetic (naked nucleic
CC acid) vaccination. The vaccines are used to prevent C. psittaci
CC infection, especially in birds.
XX
SQ Sequence 228 AA;

```
  Query Match            79.5%;  Score 58;  DB 20;  Length 228;
  Best Local Similarity  91.7%;  Pred. No. 0.029;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    182 qinkmksrkacg 193


RESULT  12
AAY56769
ID   AAY56769 standard; Protein; 343 AA.
XX
AC   AAY56769;
XX
DT   22-FEB-2000  (first entry)
XX
DE   C. trachomatis serovar GPIC MOMP sequence.
XX
KW   Major outer membrane protein; MOMP; Chlamydia; vaccine; immune response;
KW   cellular response; immunogen; Th1-like CD4 response; mucosal immunity.
XX
OS   Chlamydia trachomatis.
XX
PN   WO9951745-A2.
XX
PD   14-OCT-1999.
XX
PF   07-APR-1999;   99WO-CA00292.
XX
PR   07-APR-1998;   98US-0055765.
XX
PA   (UYMA-) UNIV MANITOBA.
XX
PI   Bruhnam RC;
XX
DR   WPI; 1999-620205/53.
XX
PT   Non-replicating vector encoding fragments of the outer membrane protein
PT   of Chlamydia, useful in vaccines and as immunogen  -
XX
PS   Disclosure; Fig 10 A-F; 52pp; English.
XX
CC   The invention provides a non-replicating vector that comprises, linked
CC   to a promoter, a nucleotide sequence that encodes a region containing at
CC   least one of the conserved domains 2, 3 and 5 of a major outer membrane
CC   protein (MOMP) of a Chlamydia strain. The vector is used: (a) in
CC   vaccines to generate a protective immune response (mainly cellular)
CC   against MOMP, and (b) as immunogens to raise anti-MOMP antibodies, useful
CC   in standard immunoassays. Immunization with the vector induces a broad
CC   spectrum of immune responses, including Th1-like CD4 responses and
CC   mucosal immunity, providing significant protection against subsequent
CC   challenge. Sequences AAY56757-71 represent MOMP sequences from a variety
CC   of serovars of C. trachomatis.
XX
SQ   Sequence   343 AA;


  Query Match            79.5%;  Score 58;  DB 20;  Length 343;
  Best Local Similarity  91.7%;  Pred. No. 0.043;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    297 qinkmksrkacg 308


RESULT  13
AAY56770
ID   AAY56770 standard; Protein; 356 AA.
XX
AC   AAY56770;
XX
DT   22-FEB-2000  (first entry)
XX
DE   C. trachomatis serovar Mn MOMP sequence.
XX
KW   Major outer membrane protein; MOMP; Chlamydia; vaccine; immune response;
KW   cellular response; immunogen; Th1-like CD4 response; mucosal immunity.
XX
OS   Chlamydia trachomatis.
XX
PN   WO9951745-A2.
XX
PD   14-OCT-1999.
XX
PF   07-APR-1999;   99WO-CA00292.
XX
PR   07-APR-1998;   98US-0055765.
XX
PA   (UYMA-) UNIV MANITOBA.
XX
PI   Bruhnam RC;
XX
DR   WPI; 1999-620205/53.
XX
PT   Non-replicating vector encoding fragments of the outer membrane protein
PT   of Chlamydia, useful in vaccines and as immunogen  -
XX
PS   Disclosure; Fig 10 A-F; 52pp; English.
XX
CC   The invention provides a non-replicating vector that comprises, linked
CC   to a promoter, a nucleotide sequence that encodes a region containing at
CC   least one of the conserved domains 2, 3 and 5 of a major outer membrane
CC   protein (MOMP) of a Chlamydia strain. The vector is used: (a) in
CC   vaccines to generate a protective immune response (mainly cellular)
CC   against MOMP, and (b) as immunogens to raise anti-MOMP antibodies, useful
CC   in standard immunoassays. Immunization with the vector induces a broad
CC   spectrum of immune responses, including Th1-like CD4 responses and
CC   mucosal immunity, providing significant protection against subsequent
CC   challenge. Sequences AAY56757-71 represent MOMP sequences from a variety
CC   of serovars of C. trachomatis.
XX
SQ   Sequence   356 AA;


  Query Match            79.5%;  Score 58;  DB 20;  Length 356;
  Best Local Similarity  91.7%;  Pred. No. 0.045;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    310 qinkmksrkacg 321


RESULT  14
AAW98188
ID   AAW98188 standard; Protein; 389 AA.
XX
AC   AAW98188;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Chlamydia psittaci major outer membrane protein.
XX
KW   Major outer membrane protein; MOMP; psittacosis; infection;
KW   vaccine; genetic immunisation.
XX
OS   Chlamydia psittaci.
XX
PN   WO9910005-A1.
XX
PD   04-MAR-1999.
XX
```

PF 28-AUG-1998; 98WO-US17943.
XX
PR 28-AUG-1997; 97US-0057147.
XX
PA (LOUU) UNIV LOUISIANA & AGRIC & MECH COLLEGE.
XX
PI Baghian A, Chouljenko VN, Kousoulas KG, Tully TN;
XX
DR WPI; 1999-254214/21.
DR N-PSDB; AAX25047.
XX
PT A new vaccine for Chlamydia psittaci infections
XX
PS Disclosure; Page 60-61; 72pp; English.
XX
CC The present sequence is the major outer membrane protein (MOMP)
CC of Chlamydia psittaci strain B577. A claimed MOMP polypeptide (see
CC AAW98184) comprises regions VD3 and VD4 of B577 MOMP, i.e. it lacks
CC regions VD1 and VD2. A claimed vaccine composition includes MOMP
CC polypeptide lacking VD1 and VD2, optionally fused to a maltose
CC binding protein. Also claimed are an isolated nucleic acid
CC encoding the polypeptide, a vector, and a method of preventing C.
CC psittaci infection by administering the vaccine containing the
CC MOMP polypeptide. Vectors encoding MOMP polypeptides lacking
CC regions VD1 and VD2 are useful for genetic vaccination. The
CC vaccines are used to prevent C. psittaci infection, especially in
CC birds.
XX
SQ Sequence 389 AA;

Query Match 79.5%; Score 58; DB 20; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.048;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QINKFKSRKACG 13
|||| |||||||
Db 343 qinkmksrkacg 354

RESULT 15
AAW98189
ID AAW98189 standard; Protein; 402 AA.
XX
AC AAW98189;
XX
DT 05-JUL-1999 (first entry)
XX
DE Chlamydia psittaci major outer membrane protein.
XX
KW Major outer membrane protein; MOMP; psittacosis; infection;
KW vaccine; genetic immunisation.
XX
OS Chlamydia psittaci.
XX
PN WO9910005-A1.
XX
PD 04-MAR-1999.
XX
PF 28-AUG-1998; 98WO-US17943.
XX
PR 28-AUG-1997; 97US-0057147.
XX
PA (LOUU) UNIV LOUISIANA & AGRIC & MECH COLLEGE.
XX
PI Baghian A, Chouljenko VN, Kousoulas KG, Tully TN;
XX
DR WPI; 1999-254214/21.
DR N-PSDB; AAX25048.
XX
PT A new vaccine for Chlamydia psittaci infections
XX
PS Disclosure; Page 65-66; 72pp; English.
XX
CC The present sequence is the major outer membrane protein (MOMP)
CC of Chlamydia psittaci strain 6BC. Claimed MOMP polypeptides (see
CC AAW98183 and AAW98184) comprise regions VD3 and VD4 of an MOMP, i.e.
CC they lack regions VD1 and VD2. Claimed vaccine compositions
CC include such MOMP polypeptides, optionally fused to a maltose
CC binding protein. Also claimed are isolated nucleic acids encoding
CC the polypeptide, a vector, and a method of preventing C. psittaci
CC infection by administering the vaccine containing the MOMP
CC polypeptide. Vectors encoding MOMP polypeptides lacking regions
CC VD1 and VD2 are useful for genetic vaccination. The vaccines are
CC used to prevent C. psittaci infection, especially in birds.
XX
SQ Sequence 402 AA;

Query Match 79.5%; Score 58; DB 20; Length 402;
Best Local Similarity 91.7%; Pred. No. 0.05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QINKFKSRKACG 13
|||| |||||||
Db 356 qinkmksrkacg 367

Search completed: March 26, 2002, 13:38:48
Job time: 142 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:41:29 ; Search time 37.72 Seconds
(without alignments)
7.756 Million cell updates/sec

Title: US-09-709-201-101
Perfect score: 73
Sequence: 1 CQINKFKSRKACG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	54.8	58	1	US-08-334-773A-2	Sequence 2, Appli
2	38	52.1	824	1	US-08-221-750A-3	Sequence 3, Appli
3	37.5	51.4	19	2	US-08-764-640-158	Sequence 158, App
4	37.5	51.4	19	3	US-08-973-225-158	Sequence 158, App
5	37.5	51.4	19	3	US-09-244-298A-158	Sequence 158, App
6	37.5	51.4	19	4	US-09-516-704-158	Sequence 158, App
7	36	49.3	42	4	US-08-974-549A-141	Sequence 141, App
8	36	49.3	562	3	US-08-851-843A-5	Sequence 5, Appli
9	36	49.3	562	4	US-08-854-050-5	Sequence 5, Appli
10	35	47.9	19	2	US-08-764-640-164	Sequence 164, App
11	35	47.9	19	3	US-08-973-225-164	Sequence 164, App
12	35	47.9	19	3	US-09-244-298A-164	Sequence 164, App
13	35	47.9	19	4	US-09-516-704-164	Sequence 164, App
14	35	47.9	68	1	US-07-972-387-40	Sequence 40, Appl
15	35	47.9	68	1	US-08-431-412-40	Sequence 40, Appl
16	35	47.9	68	1	US-08-057-971-40	Sequence 40, Appl
17	35	47.9	69	1	US-08-358-160-125	Sequence 125, App
18	35	47.9	70	1	US-07-791-213D-5	Sequence 5, Appli
19	35	47.9	70	1	US-08-293-150A-5	Sequence 5, Appli
20	35	47.9	89	1	US-07-972-387-8	Sequence 8, Appli
21	35	47.9	89	1	US-07-972-387-10	Sequence 10, Appl
22	35	47.9	89	1	US-07-972-387-16	Sequence 16, Appl
23	35	47.9	89	1	US-08-431-412-8	Sequence 8, Appli
24	35	47.9	89	1	US-08-431-412-10	Sequence 10, Appl
25	35	47.9	89	1	US-08-431-412-16	Sequence 16, Appl
26	35	47.9	89	1	US-08-057-971-8	Sequence 8, Appli
27	35	47.9	89	1	US-08-057-971-10	Sequence 10, Appl
28	35	47.9	89	1	US-08-057-971-16	Sequence 16, Appl
29	35	47.9	91	1	US-07-791-213D-89	Sequence 89, Appl
30	35	47.9	91	1	US-07-972-387-2	Sequence 2, Appli
31	35	47.9	91	1	US-08-431-412-2	Sequence 2, Appli
32	35	47.9	91	1	US-08-057-971-2	Sequence 2, Appli
33	35	47.9	91	1	US-08-293-150A-89	Sequence 89, Appl
34	35	47.9	101	1	US-07-972-387-28	Sequence 28, Appl
35	35	47.9	101	1	US-08-431-412-28	Sequence 28, Appl
36	35	47.9	101	1	US-08-057-971-28	Sequence 28, Appl
37	35	47.9	540	4	US-09-011-074-4	Sequence 4, Appli
38	35	47.9	595	4	US-08-842-079-6	Sequence 6, Appli
39	35	47.9	595	4	US-08-842-079-17	Sequence 17, Appl
40	34	46.6	58	1	US-08-334-773A-1	Sequence 1, Appli
41	34	46.6	205	1	US-08-450-944-5	Sequence 5, Appli
42	34	46.6	205	5	PCT-US96-07709-5	Sequence 5, Appli
43	34	46.6	221	1	US-08-450-944-2	Sequence 2, Appli
44	34	46.6	221	5	PCT-US96-07709-2	Sequence 2, Appli
45	34	46.6	595	4	US-08-842-079-18	Sequence 18, Appl

ALIGNMENTS

```
RESULT   1
US-08-334-773A-2
; Sequence 2, Application US/08334773A
; Patent No. 5629176
;  GENERAL INFORMATION:
;    APPLICANT:  No. 5629176ris, Fanny
;    APPLICANT:  No. 5629176ris, Kjeld
;    APPLICANT:  Bjorn, Soren Erik
;    APPLICANT:  Petersen, Lars Christian
;    APPLICANT:  Olsen, Ole Hvilsted
;    TITLE OF INVENTION:  A Human Kunitz-Type Protease Inhibitor
;    TITLE OF INVENTION:  Variant
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5629176o No. 5629176disk of No. 5629176th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 5400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6400
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/334,773A
;      FILING DATE:  04-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Agris, Cheryl H.
;      REGISTRATION NUMBER:  34,086
;      REFERENCE/DOCKET NUMBER:  3695.210-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 867 0123
;      TELEFAX:  212 878 9655
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  58 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  synthetic
US-08-334-773A-2


  Query Match            54.8%;  Score 40;  DB 1;  Length 58;
  Best Local Similarity  77.8%;  Pred. No. 5.4;
  Matches   7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      4 NKFKSRKAC 12
          |||||:| |
Db     43 NKFKSQKEC 51


RESULT   2
US-08-221-750A-3
; Sequence 3, Application US/08221750A
; Patent No. 5643747
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Steven M.
;    APPLICANT:  Deich, Robert A.
;    TITLE OF INVENTION:  Genes for the Export of Pertussis
;    TITLE OF INVENTION:  Holotoxin
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/221,750A
;      FILING DATE:  31-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/031,619
;      FILING DATE:  15-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Carroll, Alice O.
;      REGISTRATION NUMBER:  33,542
;      REFERENCE/DOCKET NUMBER:  ACC93-01A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  824 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-221-750A-3


  Query Match             52.1%;  Score 38;  DB 1;  Length 824;
  Best Local Similarity   53.8%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CQINKFKSRKACG 13
          ||: ||:|  | |
Db    471 CQLQKFRSADADG 483


RESULT   3
US-08-764-640-158
; Sequence 158, Application US/08764640
; Patent No. 5869451
; Patent No. 5869451 5837683
;  GENERAL INFORMATION:
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Gates, Christian
;    APPLICANT:  Schatz, Peter J.
;    APPLICANT:  Balasubramanian, Palaniappan
;    APPLICANT:  Wagstrom, Christopher R.
;    APPLICANT:  Hendren, Richard W.
;    APPLICANT:  Deprince, Randolph B.
;    APPLICANT:  Podduturi, Surekha
;    APPLICANT:  Yin, Qun
;    TITLE OF INVENTION:  PEPTIDES AND COMPOUNDS THAT BIND TO A
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  244
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Glaxo Wellcome
;      STREET:  Five Moore Drive, P.O. Box 13398
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,640
;      FILING DATE:  11-DEC-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hrubiec, Robert T.
;      REGISTRATION NUMBER:  36,392
;      REFERENCE/DOCKET NUMBER:  PK3281
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-248-1000
;  INFORMATION FOR SEQ ID NO:  158:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-764-640-158


  Query Match             51.4%;  Score 37.5;  DB 2;  Length 19;
  Best Local Similarity   61.5%;  Pred. No. 4.9;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy      1 CQINKFKSRKACG 13
          | :| ||||  ||
Db      5 CTLNGFKSRH-CG 16


RESULT   4
US-08-973-225-158
; Sequence 158, Application US/08973225A
; Patent No. 6083913
;   GENERAL INFORMATION:
;        APPLICANT: Dower, William J.
;                   Barrett, Ronald W.
;                   Cwirla, Steven E.
;                   Duffin, David J.
;                   Gates, Christian
;                   Haselden, Sherril S.
;                   Mattheakis, Larry C.
;                   Schatz, Peter J.
;                   Wagstrom, Christopher R.
;                   Wrighton, Nicholas C.
;        TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                            THROMBOPOIETIN RECEPTOR
;        NUMBER OF SEQUENCES: 232
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Glaxo Wellcome
;             STREET: Five Moore Drive, P.O. Box 13398
;             CITY: Research Triangle Park
;             STATE: NC
;             COUNTRY: USA
;             ZIP: 27709
```

```
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/973,225A
;             FILING DATE: 04-Dec-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hrubiec, Robert T.
;             REGISTRATION NUMBER: 36,392
;             REFERENCE/DOCKET NUMBER: PK3065USW
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 919-248-1000
;   INFORMATION FOR SEQ ID NO: 158:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 19 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 158:
US-08-973-225-158


  Query Match             51.4%;  Score 37.5;  DB 3;  Length 19;
  Best Local Similarity   61.5%;  Pred. No. 4.9;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy      1 CQINKFKSRKACG 13
          | :| ||||  ||
Db      5 CTLNGFKSRH-CG 16


RESULT   5
US-09-244-298A-158
; Sequence 158, Application US/09244298A
; Patent No. 6121238
;  GENERAL INFORMATION:
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Gates, Christian
;    APPLICANT:  Schatz, Peter J.
;    APPLICANT:  Balasubramanian, Palaniappan
;    APPLICANT:  Wagstrom, Christopher R.
;    APPLICANT:  Hendren, Richard W.
;    APPLICANT:  Deprince, Randolph B.
;    APPLICANT:  Podduturi, Surekha
;    APPLICANT:  Yin, Qun
;    TITLE OF INVENTION:  PEPTIDES AND COMPOUNDS THAT BIND TO A
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  244
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Glaxo Wellcome
;      STREET:  Five Moore Drive, P.O. Box 13398
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/244,298A
;      FILING DATE:  11-DEC-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hrubiec, Robert T.
;      REGISTRATION NUMBER:  36,392
;      REFERENCE/DOCKET NUMBER:  PK3281
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-248-1000
;  INFORMATION FOR SEQ ID NO:  158:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-244-298A-158


  Query Match             51.4%;  Score 37.5;  DB 3;  Length 19;
  Best Local Similarity   61.5%;  Pred. No. 4.9;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy      1 CQINKFKSRKACG 13
          | :| ||||  ||
Db      5 CTLNGFKSRH-CG 16


RESULT   6
US-09-516-704-158
; Sequence 158, Application US/09516704
; Patent No. 6251864
;  GENERAL INFORMATION:
;       APPLICANT: Dower, William J.
;                  Barrett, Ronald W.
;                  Cwirla, Steven E.
;                  Gates, Christian
;                  Schatz, Peter J.
;                  Balasubramanian, Palaniappan
;                  Wagstrom, Christopher R.
;                  Hendren, Richard W.
;                  Deprince, Randolph B.
;                  Podduturi, Surekha
;       TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                           RECEPTOR
;       NUMBER OF SEQUENCES: 244
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Glaxo Wellcome
;            STREET: Five Moore Drive, P.O. Box 13398
;            CITY: Research Triangle Park
;            STATE: NC
;            COUNTRY: USA
;            ZIP: 27709
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/516,704
;            FILING DATE: 01-Mar-2000
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hrubiec, Robert T.
;            REGISTRATION NUMBER: 36,392
;            REFERENCE/DOCKET NUMBER: PK3281
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 919-248-1000
;  INFORMATION FOR SEQ ID NO: 158:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 19 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 158:
US-09-516-704-158


  Query Match             51.4%;  Score 37.5;  DB 4;  Length 19;
```

```
  Best Local Similarity   61.5%;  Pred. No. 4.9;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       1 CQINKFKSRKACG 13
           | :| ||||  ||
Db       5 CTLNGFKSRH-CG 16


RESULT   7
US-08-974-549A-141
; Sequence 141, Application US/08974549A
; Patent No. 6166178
;  GENERAL INFORMATION:
;    APPLICANT:  Cech, Thomas R.
;    APPLICANT:  Lingner, Joachim
;    APPLICANT:  Nakamura, Toru
;    APPLICANT:  Chapman, Karen B.
;    APPLICANT:  Morin, Gregg B.
;    APPLICANT:  Harley, Calvin B.
;    APPLICANT:  Andrews, William H.
;    TITLE OF INVENTION:  Human Telomerase Catalytic Subunit
;    NUMBER OF SEQUENCES:  727
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,549A
;      FILING DATE:  19-NOV-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/724,643
;      FILING DATE:  01-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/844,419
;      FILING DATE:  18-APR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/846,017
;      FILING DATE:  25-APR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/851,843
;      FILING DATE:  06-MAY-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/854,050
;      FILING DATE:  09-MAY-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/911,312
;      FILING DATE:  14-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/912,951
;      FILING DATE:  14-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/915,503
;      FILING DATE:  14-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US97/17618
;      FILING DATE:  01-OCT-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US97/17885
;      FILING DATE:  01-OCT-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Apple, Randolph Ted
;      REGISTRATION NUMBER:  36,429
;      REFERENCE/DOCKET NUMBER:  015389-002610US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  141:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-974-549A-141


  Query Match           49.3%;  Score 36;  DB 4;  Length 42;
  Best Local Similarity   60.0%;  Pred. No. 17;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 NKFKSRKACG 13
           | :| || ||
Db       8 NPYKKRKNCG 17


RESULT   8
US-08-851-843A-5
; Sequence 5, Application US/08851843A
; Patent No. 6093809
;  GENERAL INFORMATION:
;    APPLICANT:  Cech, Thomas R.
;    APPLICANT:  Lingner, Joachim
;    APPLICANT:  Nakamura, Toru
;    APPLICANT:  Chapman, Karen B.
;    APPLICANT:  Morin, Gregg B.
;    APPLICANT:  Harley, Calvin
;    APPLICANT:  Andrews, William H.
;    TITLE OF INVENTION:  No. 6093809el Telomerase
;    NUMBER OF SEQUENCES:  225
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/851,843A
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/846,017
;      FILING DATE:  25-APR-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/844,419
;      FILING DATE:  18-APR-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/724,643
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Apple,  Randolph T.
;      REGISTRATION NUMBER:  36,429
;      REFERENCE/DOCKET NUMBER:  015389-002930US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  562 amino acids
```

```
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-851-843A-5


  Query Match             49.3%;  Score 36;  DB 3;  Length 562;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 NKFKSRKACG 13
          | :| || ||
Db      8 NPYKKRKNCG 17


RESULT   9
US-08-854-050-5
; Sequence 5, Application US/08854050
; Patent No. 6261836
;  GENERAL INFORMATION:
;    APPLICANT:  Cech, Thomas R.
;    APPLICANT:  Lingner, Joachim
;    APPLICANT:  Nakamura, Toru
;    APPLICANT:  Chapman, Karen B.
;    APPLICANT:  Morin, Gregg B.
;    APPLICANT:  Harley, Calvin
;    APPLICANT:  Andrews, William H.
;    TITLE OF INVENTION:  No. 6261836el Telomerase
;    NUMBER OF SEQUENCES:  225
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/854,050
;      FILING DATE:  09-MAY-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/851,843
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/846,017
;      FILING DATE:  25-APR-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/844,419
;      FILING DATE:  18-APR-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/724,643
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Apple,  Randolph T.
;      REGISTRATION NUMBER:  36,429
;      REFERENCE/DOCKET NUMBER:  015389-002930US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  562 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-854-050-5


  Query Match             49.3%;  Score 36;  DB 4;  Length 562;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 NKFKSRKACG 13
          | :| || ||
Db      8 NPYKKRKNCG 17


RESULT  10
US-08-764-640-164
; Sequence 164, Application US/08764640
; Patent No. 5869451
; Patent No. 5869451 5837683
;  GENERAL INFORMATION:
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Gates, Christian
;    APPLICANT:  Schatz, Peter J.
;    APPLICANT:  Balasubramanian, Palaniappan
;    APPLICANT:  Wagstrom, Christopher R.
;    APPLICANT:  Hendren, Richard W.
;    APPLICANT:  Deprince, Randolph B.
;    APPLICANT:  Podduturi, Surekha
;    APPLICANT:  Yin, Qun
;    TITLE OF INVENTION:  PEPTIDES AND COMPOUNDS THAT BIND TO A
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  244
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Glaxo Wellcome
;      STREET:  Five Moore Drive, P.O. Box 13398
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,640
;      FILING DATE:  11-DEC-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hrubiec, Robert T.
;      REGISTRATION NUMBER:  36,392
;      REFERENCE/DOCKET NUMBER:  PK3281
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-248-1000
;  INFORMATION FOR SEQ ID NO:  164:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-764-640-164


  Query Match             47.9%;  Score 35;  DB 2;  Length 19;
  Best Local Similarity   46.2%;  Pred. No. 12;
  Matches    6;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 CQINKFKSRKACG 13
          | : | |   |||
```

```
Db      5 CSLAKLKKGAACG 17


RESULT  11
US-08-973-225-164
; Sequence 164, Application US/08973225A
; Patent No. 6083913
;   GENERAL INFORMATION:
;        APPLICANT: Dower, William J.
;                   Barrett, Ronald W.
;                   Cwirla, Steven E.
;                   Duffin, David J.
;                   Gates, Christian
;                   Haselden, Sherril S.
;                   Mattheakis, Larry C.
;                   Schatz, Peter J.
;                   Wagstrom, Christopher R.
;                   Wrighton, Nicholas C.
;        TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                            THROMBOPOIETIN RECEPTOR
;        NUMBER OF SEQUENCES: 232
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Glaxo Wellcome
;             STREET: Five Moore Drive, P.O. Box 13398
;             CITY: Research Triangle Park
;             STATE: NC
;             COUNTRY: USA
;             ZIP: 27709
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/973,225A
;             FILING DATE: 04-Dec-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hrubiec, Robert T.
;             REGISTRATION NUMBER: 36,392
;             REFERENCE/DOCKET NUMBER: PK3065USW
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 919-248-1000
;   INFORMATION FOR SEQ ID NO: 164:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 19 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 164:
US-08-973-225-164


  Query Match             47.9%;  Score 35;  DB 3;  Length 19;
  Best Local Similarity   46.2%;  Pred. No. 12;
  Matches    6;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 CQINKFKSRKACG 13
          | : | |   |||
Db      5 CSLAKLKKGAACG 17


RESULT  12
US-09-244-298A-164
; Sequence 164, Application US/09244298A
; Patent No. 6121238
;  GENERAL INFORMATION:
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Gates, Christian
;    APPLICANT:  Schatz, Peter J.
;    APPLICANT:  Balasubramanian, Palaniappan
;    APPLICANT:  Wagstrom, Christopher R.
;    APPLICANT:  Hendren, Richard W.
;    APPLICANT:  Deprince, Randolph B.
;    APPLICANT:  Podduturi, Surekha
;    APPLICANT:  Yin, Qun
;    TITLE OF INVENTION:  PEPTIDES AND COMPOUNDS THAT BIND TO A
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  244
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Glaxo Wellcome
;      STREET:  Five Moore Drive, P.O. Box 13398
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/244,298A
;      FILING DATE:  11-DEC-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hrubiec, Robert T.
;      REGISTRATION NUMBER:  36,392
;      REFERENCE/DOCKET NUMBER:  PK3281
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-248-1000
;  INFORMATION FOR SEQ ID NO:  164:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-244-298A-164


  Query Match             47.9%;  Score 35;  DB 3;  Length 19;
  Best Local Similarity   46.2%;  Pred. No. 12;
  Matches    6;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 CQINKFKSRKACG 13
          | : | |   |||
Db      5 CSLAKLKKGAACG 17


RESULT  13
US-09-516-704-164
; Sequence 164, Application US/09516704
; Patent No. 6251864
;   GENERAL INFORMATION:
;        APPLICANT: Dower, William J.
;                   Barrett, Ronald W.
;                   Cwirla, Steven E.
;                   Gates, Christian
;                   Schatz, Peter J.
;                   Balasubramanian, Palaniappan
;                   Wagstrom, Christopher R.
;                   Hendren, Richard W.
;                   Deprince, Randolph B.
;                   Podduturi, Surekha
;        TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                            RECEPTOR
;        NUMBER OF SEQUENCES: 244
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Glaxo Wellcome
;             STREET: Five Moore Drive, P.O. Box 13398
;             CITY: Research Triangle Park
;             STATE: NC
;             COUNTRY: USA
```

```
;            ZIP: 27709
;        COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/516,704
;            FILING DATE: 01-Mar-2000
;            CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;            NAME: Hrubiec, Robert T.
;            REGISTRATION NUMBER: 36,392
;            REFERENCE/DOCKET NUMBER: PK3281
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 919-248-1000
;    INFORMATION FOR SEQ ID NO: 164:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 19 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 164:
US-09-516-704-164


  Query Match             47.9%;  Score 35;  DB 4;  Length 19;
  Best Local Similarity   46.2%;  Pred. No. 12;
  Matches    6;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 CQINKFKSRKACG 13
           | : | |   |||
Db       5 CSLAKLKKGAACG 17


RESULT  14
US-07-972-387-40
; Sequence 40, Application US/07972387
; Patent No. 5451659
;  GENERAL INFORMATION:
;    APPLICANT:  Morishita, Hideaki
;    APPLICANT:  Kanamori, Toshinori
;    APPLICANT:  No. 5451659uhara, Masahiro
;    TITLE OF INVENTION:  Polypeptide, DNA Fragment Encoding the
;    TITLE OF INVENTION:  Same, Drug Composition Containing the Same and Process for
;    TITLE OF INVENTION:  Producing the Same
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  301 N. Washington St.
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/972,387
;      FILING DATE:  19921105
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1110-124P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  68 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  C-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Eschericia coli
US-07-972-387-40


  Query Match             47.9%;  Score 35;  DB 1;  Length 68;
  Best Local Similarity   66.7%;  Pred. No. 39;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 NKFKSRKAC 12
           |||:| | |
Db      41 NKFESEKEC 49


RESULT  15
US-08-431-412-40
; Sequence 40, Application US/08431412
; Patent No. 5589360
;  GENERAL INFORMATION:
;    APPLICANT:  Morishita, Hideaki
;    APPLICANT:  Kanamori, Toshinori
;    APPLICANT:  No. 5589360uhara, Masahiro
;    TITLE OF INVENTION:  Polypeptide, DNA Fragment Encoding the
;    TITLE OF INVENTION:  Same, Drug Composition Containing the Same and Process for
;    TITLE OF INVENTION:  Producing the Same
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  301 N. Washington St.
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/431,412
;      FILING DATE:  28-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/972,387
;      FILING DATE:  05-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1110-124P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  68 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  C-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Eschericia coli
US-08-431-412-40
```

Query Match 47.9%; Score 35; DB 1; Length 68;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 NKFKSRKAC 12
           |||:| | |
Db      41 NKFESEKEC 49
```

Search completed: March 26, 2002, 13:41:30
Job time: 304 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:37:21 ; Search time 42.75 Seconds
(without alignments)
23.164 Million cell updates/sec

Title: US-09-709-201-101
Perfect score: 73
Sequence: 1 CQINKFKSRKACG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73	100.0	389	2	A43587	major outer membra
2	73	100.0	389	2	I40864	major outer membra
3	73	100.0	389	2	I40739	major outer membra
4	73	100.0	389	2	D86577	major outer membra
5	58	79.5	389	1	MMCWP3	major outer membra
6	58	79.5	389	2	A60109	major outer membra
7	58	79.5	392	2	A40371	major outer membra
8	58	79.5	402	1	MMCWPM	major outer membra
9	58	79.5	402	2	A60341	major outer membra
10	58	79.5	402	2	B60109	major outer membra
11	58	79.5	402	2	I40740	major outer membra
12	53	72.6	372	2	S11009	major outer membra
13	53	72.6	372	2	B60756	major outer membra
14	53	72.6	374	2	S11006	major outer membra
15	53	72.6	375	2	S11007	major outer membra
16	53	72.6	387	2	JT0947	mouse pneumonitis
17	53	72.6	387	2	S16034	major outer membra
18	53	72.6	387	2	C81747	major outer membra
19	53	72.6	393	1	MMCWTE	major outer membra
20	53	72.6	393	2	T01645	major outer membra
21	53	72.6	393	2	S06259	major outer membra
22	53	72.6	393	2	JC1432	major outer membra
23	53	72.6	393	2	H71484	probable major out
24	53	72.6	394	1	MMCWTB	major outer membra
25	53	72.6	394	2	S11012	major outer membra
26	53	72.6	395	1	MMCWTF	major outer membra
27	53	72.6	396	2	S12799	major outer membra
28	53	72.6	397	1	MMCWTH	major outer membra
29	53	72.6	397	1	MMCWTC	major outer membra
30	53	72.6	397	2	JE0413	major outer membra
31	53	72.6	404	2	I40741	major outer membra
32	43	58.9	265	2	D81402	hypothetical prote
33	41	56.2	1154	2	T18525	diacylglycerol kin
34	40	54.8	182	2	T28390	ORF MSV229 leucine
35	39	53.4	862	1	A49346	aldehyde dehydroge
36	39	53.4	894	2	D82127	alcohol dehydrogen
37	38	52.1	444	2	D84948	NADH dehydrogenase
38	38	52.1	824	2	B47301	VirB4 homolog - Bo
39	38	52.1	1330	2	A36373	hypothetical prote
40	38	52.1	1333	2	T38401	retrotransposable
41	37	50.7	235	2	T44466	transposase tnpA [
42	37	50.7	364	2	S43117	transposase - Lept
43	37	50.7	372	2	JC2556	alpha-1-microglobu
44	37	50.7	408	2	T16601	hypothetical prote
45	37	50.7	444	2	H82821	NADH-ubiquinone ox

ALIGNMENTS

RESULT 1
A43587
major outer membrane protein, porin CP0051 precursor [imported] - Chlamydophila pneum
N;Alternate names: MOMP
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 11-May-2000
C;Accession: A43587; A49751; A49216; G72044; F81619
R;Perez Melgosa, M.; Kuo, C.C.; Campbell, L.A.
Infect. Immun. 59, 2195-2199, 1991
A;Title: Sequence analysis of the major outer membrane protein gene of Chlamydia pneu
A;Reference number: A43587; MUID:91244474
A;Accession: A43587
A;Molecule type: DNA
A;Residues: 1-389 <PER>
A;Cross-references: GB:M69230; NID:g144540; PIDN:AAA73071.1; PID:g144541
R;Carter, M.W.; Al-Mahdawi, S.A.H.; Giles, I.G.; Treharne, J.D.; Ward, M.E.; Clarke,
J. Gen. Microbiol. 137, 465-475, 1991
A;Title: Nucleotide sequence and taxonomic value of the major outer membrane protein
A;Reference number: A49751; MUID:91237311
A;Accession: A49751
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <CAR>
A;Cross-references: GB:M64064; GB:M34942; NID:g144534; PIDN:AAA23143.1; PID:g144535
A;Note: isolate IOL-207
R;Gaydos, C.A.; Quinn, T.C.; Bobo, L.D.; Eiden, J.J.
Infect. Immun. 60, 5319-5323, 1992
A;Title: Similarity of Chlamydia pneumoniae strains in the variable domain IV region
A;Reference number: A49216; MUID:93084388
A;Accession: A49216
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 297-352 <GAY>
A;Cross-references: GB:S50607; NID:g260972; PIDN:AAB24363.1; PID:g260973
A;Note: sequence extracted from NCBI backbone (NCBIN:120604, NCBIP:120605)
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: G72044
A;Molecule type: DNA
A;Residues: 1-389 <ARN>
A;Cross-references: GB:AE001652; GB:AE001363; NID:g4376997; PIDN:AAD18834.1; PID:g437
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255
A;Accession: F81619
A;Status: preliminary
A;Molecule type: DNA

```
A;Residues: 1-389 <REA>
A;Cross-references: GB:AE002168; GB:AE002161; NID:g7188982; PIDN:AAF37944.1; PID:g718899
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: ompA; CP0051
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match             100.0%;  Score 73;  DB 2;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CQINKFKSRKACG 13
           |||||||||||||
Db     342 CQINKFKSRKACG 354


RESULT   2
I40864
major outer membrane protein - Chlamydophila psittaci
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
C;Accession: I40864; S33465
R;Girjes, A.A.; Carrick, F.N.; Lavin, M.F.
Gene 138, 139-142, 1994
A;Title: Remarkable sequence relatedness in the DNA encoding the major outer membrane pr
A;Reference number: I40864; MUID:94171025
A;Accession: I40864
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-389 <RES>
A;Cross-references: EMBL:X72023; NID:g313844; PIDN:CAA50906.1; PID:g313845
C;Superfamily: Chlamydia major outer membrane protein


  Query Match             100.0%;  Score 73;  DB 2;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CQINKFKSRKACG 13
           |||||||||||||
Db     342 CQINKFKSRKACG 354


RESULT   3
I40739
major outer membrane protein precursor - Chlamydophila pneumoniae (strain equine/N16)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
A;Variety: strain equine/N16
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Apr-2000
C;Accession: I40739
R;Storey, C.; Lusher, M.; Yates, P.; Richmond, S.
J. Gen. Microbiol. 139, 2621-2626, 1993
A;Title: Evidence for Chlamydia pneumoniae of non-human origin.
A;Reference number: I40739; MUID:94103736
A;Accession: I40739
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-389 <STO>
A;Cross-references: GB:L04982; NID:g289840; PIDN:AAA17397.1; PID:g289841
C;Comment: On the basis of the major outer membrane protein the authors classified the e
 the sequence of the genome strain CWL029 and strain strain IOL-207. See PIR:A43587.
C;Genetics:
A;Gene: momp
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match             100.0%;  Score 73;  DB 2;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CQINKFKSRKACG 13
           |||||||||||||
Db     342 CQINKFKSRKACG 354


RESULT   4
D86577
major outer membrane protein [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: D86577
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: D86577
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <STO>
A;Cross-references: GB:BA000008; NID:g8979067; PIDN:BAA98902.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: ompA
C;Superfamily: Chlamydia major outer membrane protein


  Query Match             100.0%;  Score 73;  DB 2;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CQINKFKSRKACG 13
           |||||||||||||
Db     342 CQINKFKSRKACG 354


RESULT   5
MMCWP3
major outer membrane protein precursor - Chlamydophila psittaci (strain S26/3)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 31-Mar-2000
C;Accession: S08770
R;Herring, A.J.; Tan, T.W.; Baxter, S.; Inglis, N.F.; Dunbar, S.
FEMS Microbiol. Lett. 65, 153-158, 1989
A;Title: Sequence analysis of the major outer membrane protein gene of an ovine abort
A;Reference number: S08770
A;Accession: S08770
A;Molecule type: DNA
A;Residues: 1-389 <HER>
A;Cross-references: EMBL:X51859; NID:g40600; PIDN:CAA36152.1; PID:g40601
C;Superfamily: Chlamydia major outer membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-389/Product: major outer membrane protein #status predicted <MAT>


  Query Match             79.5%;  Score 58;  DB 1;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.011;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     343 QINKMKSRKACG 354


RESULT   6
A60109
major outer membrane protein precursor - Chlamydophila psittaci (strain Guinea pig in
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 31-Mar-2000
C;Accession: A60109
```

R;Zhang, Y.X.; Morrison, S.G.; Caldwell, H.D.; Baehr, W.
Infect. Immun. 57, 1621-1625, 1989
A;Title: Cloning and sequence analysis of the major outer membrane protein genes of two
A;Reference number: A60109; MUID:89212917
A;Accession: A60109
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-389 <ZHA>
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-389/Product: major outer membrane protein #status predicted <MAT>

Query Match 79.5%; Score 58; DB 2; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QINKFKSRKACG 13
|||| |||||||
Db 343 QINKMKSRKACG 354

RESULT 7
A40371
major outer membrane protein precursor - Chlamydophila psittaci (strain Fpn/pring)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 31-Mar-2000
C;Accession: I40859; A40371; S16137
R;Storey, C.; Lusher, M.; Yates, P.; Richmond, S.
J. Gen. Microbiol. 139, 2621-2626, 1993
A;Title: Evidence for Chlamydia pneumoniae of non-human origin.
A;Reference number: I40739; MUID:94103736
A;Accession: I40859
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMB
A;Molecule type: DNA
A;Residues: 1-392 <RES>
A;Cross-references: EMBL:X61096; NID:g40564; PIDN:CAA43409.1; PID:g40565
A;Experimental source: strain Fpn
C;Genetics:
A;Gene: MOMP
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-392/Product: major outer membrane protein #status predicted <MAT>

Query Match 79.5%; Score 58; DB 2; Length 392;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QINKFKSRKACG 13
|||| |||||||
Db 346 QINKMKSRKACG 357

RESULT 8
MMCWPM
major outer membrane protein precursor - Chlamydophila psittaci (strain A22/M)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 31-Mar-2000
C;Accession: S05954
R;Pickett, M.A.; Everson, J.S.; Clarke, I.N.
FEMS Microbiol. Lett. 55, 229-234, 1988
A;Title: Chlamydia psittaci ewe abortion agent: complete nucleotide sequence of the majc
A;Reference number: S05954
A;Accession: S05954
A;Molecule type: DNA
A;Residues: 1-402 <PIC>
A;Cross-references: EMBL:X12647; NID:g40604; PIDN:CAA31177.1; PID:g40605
C;Superfamily: Chlamydia major outer membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-402/Product: major outer membrane protein #status predicted <MAT>

Query Match 79.5%; Score 58; DB 1; Length 402;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QINKFKSRKACG 13
|||| |||||||
Db 356 QINKMKSRKACG 367

RESULT 9
A60341
major outer membrane protein precursor - Chlamydophila psittaci (strain 6BC)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 11-Dec-1992 #sequence_revision 24-Feb-1994 #text_change 31-Mar-2000
C;Accession: A44565; A60341; B60341
R;Everett, K.D.E.
submitted to the EMBL Data Library, December 1990
A;Reference number: A44565
A;Accession: A44565
A;Molecule type: DNA
A;Residues: 1-402 <EVE>
A;Cross-references: GB:X56980; NID:g40568; PIDN:CAA40300.1; PID:g40569
R;Everett, K.D.E.; Andersen, A.A.; Plaunt, M.; Hatch, T.P.
Infect. Immun. 59, 2853-2855, 1991
A;Title: Cloning and sequence analysis of the major outer membrane protein gene of Ch
A;Reference number: A60341; MUID:91310346
A;Accession: A60341
A;Molecule type: protein
A;Residues: 23-35 <EV2>
A;Accession: B60341
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 112-232;317-349 <EV3>
A;Cross-references: GB:X56980
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-402/Product: major outer membrane protein #status experimental <MAT>

Query Match 79.5%; Score 58; DB 2; Length 402;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QINKFKSRKACG 13
|||| |||||||
Db 356 QINKMKSRKACG 367

RESULT 10
B60109
major outer membrane protein precursor - Chlamydophila psittaci (strain meningopneumo
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 31-Mar-2000
C;Accession: B60109
R;Zhang, Y.X.; Morrison, S.G.; Caldwell, H.D.; Baehr, W.
Infect. Immun. 57, 1621-1625, 1989
A;Title: Cloning and sequence analysis of the major outer membrane protein genes of t
A;Reference number: A60109; MUID:89212917
A;Accession: B60109
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-402 <ZHA>
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-389/Product: major outer membrane protein #status predicted <MAT>

Query Match 79.5%; Score 58; DB 2; Length 402;
Best Local Similarity 91.7%; Pred. No. 0.011;

```
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    356 QINKMKSRKACG 367


RESULT  11
I40740
major outer membrane protein - Chlamydophila psittaci (strain N352)
C;Species: Chlamydophila psittaci, Chlamydia psittaci
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
C;Accession: I40740
R;Storey, C.; Lusher, M.; Yates, P.; Richmond, S.
J. Gen. Microbiol. 139, 2621-2626, 1993
A;Title: Evidence for Chlamydia pneumoniae of non-human origin.
A;Reference number: I40739; MUID:94103736
A;Accession: I40740
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EME
A;Molecule type: DNA
A;Residues: 1-402 <RES>
A;Cross-references: GB:L04980; NID:g144544; PIDN:AAA17396.1; PID:g144545
C;Genetics:
A;Gene: momp
C;Superfamily: Chlamydia major outer membrane protein


  Query Match             79.5%;  Score 58;  DB 2;  Length 402;
  Best Local Similarity   91.7%;  Pred. No. 0.011;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    356 QINKMKSRKACG 367


RESULT  12
S11009
major outer membrane protein - Chlamydia trachomatis (serotype B)
C;Species: Chlamydia trachomatis
A;Variety: serotype B
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 30-Apr-1999
C;Accession: S11009
R;Baehr, W.; Zhang, Y.X.; Joseph, T.; Su, H.; Nano, F.E.; Everett, K.D.E.; Caldwell, H.D
Proc. Natl. Acad. Sci. U.S.A. 85, 4000-4004, 1988
A;Title: Mapping antigenic domains expressed by Chlamydia trachomatis major outer membra
A;Reference number: S11006; MUID:88234546
A;Accession: S11009
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-372 <BAE>
A;Experimental source: serovar B
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-372/Product: major outer membrane protein #status predicted <MAT>


  Query Match             72.6%;  Score 53;  DB 2;  Length 372;
  Best Local Similarity   75.0%;  Pred. No. 0.082;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| ||||:||
Db    326 QLNKMKSRKSCG 337


RESULT  13
B60756
major outer membrane protein - Chlamydia trachomatis (serotype B)
C;Species: Chlamydia trachomatis
C;Date: 05-Jun-1993 #sequence_revision 24-Feb-1994 #text_change 07-May-1999
C;Accession: B60756
R;Hayes, L.J.; Pickett, M.A.; Conlan, J.W.; Ferris, S.; Everson, J.S.; Ward, M.E.; Clark
J. Gen. Microbiol. 136, 1559-1566, 1990
A;Title: The major outer-membrane proteins of Chlamydia trachomatis serovars A and B:
g domains.
A;Reference number: A60756; MUID:91086917
A;Accession: B60756
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-372 <HAY>
A;Experimental source: strain B/Jali-20/OT
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein


  Query Match             72.6%;  Score 53;  DB 2;  Length 372;
  Best Local Similarity   75.0%;  Pred. No. 0.082;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| ||||:||
Db    326 QLNKMKSRKSCG 337


RESULT  14
S11006
major outer membrane protein - Chlamydia trachomatis (serotype A)
C;Species: Chlamydia trachomatis
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 18-Jun-1993
C;Accession: S11006
R;Baehr, W.; Zhang, Y.X.; Joseph, T.; Su, H.; Nano, F.E.; Everett, K.D.E.; Caldwell,
Proc. Natl. Acad. Sci. U.S.A. 85, 4000-4004, 1988
A;Title: Mapping antigenic domains expressed by Chlamydia trachomatis major outer mem
A;Reference number: S11006; MUID:88234546
A;Accession: S11006
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-374 <BAE>
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-374/Product: major outer membrane protein #status predicted <MAT>


  Query Match             72.6%;  Score 53;  DB 2;  Length 374;
  Best Local Similarity   75.0%;  Pred. No. 0.082;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| ||||:||
Db    328 QLNKMKSRKSCG 339


RESULT  15
S11007
major outer membrane protein - Chlamydia trachomatis (serotype C)
C;Species: Chlamydia trachomatis
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 18-Jun-1993
C;Accession: S11007
R;Baehr, W.; Zhang, Y.X.; Joseph, T.; Su, H.; Nano, F.E.; Everett, K.D.E.; Caldwell,
Proc. Natl. Acad. Sci. U.S.A. 85, 4000-4004, 1988
A;Title: Mapping antigenic domains expressed by Chlamydia trachomatis major outer mem
A;Reference number: S11006; MUID:88234546
A;Accession: S11007
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-375 <BAE>
C;Superfamily: Chlamydia major outer membrane protein
C;Keywords: membrane protein
F;1-375/Product: major outer membrane protein #status predicted <MAT>


  Query Match             72.6%;  Score 53;  DB 2;  Length 375;
  Best Local Similarity   75.0%;  Pred. No. 0.082;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       2 QINKFKSRKACG 13
           |:|| ||||:||
Db     329 QLNKMKSRKSCG 340
```

Search completed: March 26, 2002, 13:37:21
Job time: 55 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:40:45 ; Search time 24.63 Seconds
(without alignments)
19.352 Million cell updates/sec

Title: US-09-709-201-101
Perfect score: 73
Sequence: 1 CQINKFKSRKACG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73	100.0	333	1	OM1K_CHLPN	Q9xbf4 chlamydia p
2	73	100.0	389	1	OM1N_CHLPN	Q07430 chlamydia p
3	73	100.0	389	1	OMP1_CHLPN	P27455 chlamydia p
4	58	79.5	389	1	OM1A_CHLPS	P16567 chlamydia p
5	58	79.5	392	1	OM1P_CHLPS	Q00087 chlamydia p
6	58	79.5	402	1	OM1E_CHLPS	P10332 chlamydia p
7	53	72.6	387	1	OMP1_CHLMU	P75024 chlamydia m
8	53	72.6	393	1	OM1D_CHLTR	Q46409 chlamydia t
9	53	72.6	393	1	OM1E_CHLTR	P17451 chlamydia t
10	53	72.6	393	1	OM1L_CHLTR	P19542 chlamydia t
11	53	72.6	394	1	OM1B_CHLTR	P23421 chlamydia t
12	53	72.6	394	1	OM1M_CHLTR	P06597 chlamydia t
13	53	72.6	395	1	OM1F_CHLTR	P16155 chlamydia t
14	53	72.6	396	1	OM1A_CHLTR	P23732 chlamydia t
15	53	72.6	397	1	OM1C_CHLTR	P08780 chlamydia t
16	53	72.6	397	1	OM1H_CHLTR	P13467 chlamydia t
17	53	72.6	397	1	OM1N_CHLTR	P23114 chlamydia t
18	41	56.2	1154	1	KDGD_MESAU	Q64398 mesocricetu
19	41	56.2	1195	1	KDGD_HUMAN	Q16760 homo sapien
20	39	53.4	862	1	ADHE_CLOAB	P33744 clostridium
21	38	52.1	444	1	NUOF_BUCAI	P57256 buchnera ap
22	38	52.1	1333	1	RDPO_SCHPO	Q05654 schizosacch
23	37	50.7	1075	1	PST2_SCHPO	O13919 schizosacch
24	37	50.7	1419	1	LYS2_SCHPO	P40976 schizosacch
25	37	50.7	1507	1	CADE_DROME	Q24298 drosophila
26	36	49.3	165	1	YSEA_STACA	P47995 staphylococ
27	36	49.3	419	1	CG2B_ORYSA	Q40671 oryza sativ
28	35.5	48.6	139	1	RS12_DROME	P80455 drosophila
29	35	47.9	141	1	LCA_PIG	P18137 sus scrofa
30	35	47.9	172	1	R172_HORVU	P35267 hordeum vul
31	35	47.9	300	1	TFPI_RABIT	P19761 oryctolagus
32	35	47.9	331	1	CATV_NPVBS	Q9ywk4 buzura supp
33	35	47.9	557	1	TKT2_HUMAN	P51854 homo sapien
34	35	47.9	595	1	P2X7_MOUSE	Q9z1m0 mus musculu
35	35	47.9	595	1	P2X7_RAT	Q64663 rattus norv
36	35	47.9	704	1	TRFL_PIG	P14632 sus scrofa
37	35	47.9	884	1	RPOL_BPT3	P07659 bacteriopha
38	35	47.9	2264	1	POL1_TBRVS	P18522 tomato blac
39	34.5	47.3	373	1	EGON_DROME	P15370 drosophila
40	34.5	47.3	512	1	DNB2_ADE04	P06500 human adeno
41	34.5	47.3	517	1	DNB2_ADE07	P04497 human adeno
42	34	46.6	36	1	TXAM_METSE	P11495 metridium s
43	34	46.6	146	1	AP4A_HUMAN	P50583 homo sapien
44	34	46.6	146	1	AP4A_MOUSE	P56380 mus musculu
45	34	46.6	146	1	AP4A_PIG	P50584 sus scrofa

ALIGNMENTS

RESULT 1
OM1K_CHLPN
ID OM1K_CHLPN STANDARD; PRT; 333 AA.
AC Q9XBF4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN (MOMP) (FRAGMENT).
GN OMPA OR OMP1.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KOALA TYPE I;
RX MEDLINE=93123168; PubMed=8419295;
RA Kaltenboeck B., Kousoulas K.G., Storz J.;
RT "Structures of and allelic diversity and relationships among the major
RT outer membrane protein (ompA) genes of the four chlamydial species.";
RL J. Bacteriol. 175:487-502(1993).
CC -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC ---

CC ---
DR EMBL; M73038; AAD38210.1; -.
DR InterPro; IPR000604; Chlamydia_OMP.
DR Pfam; PF01308; Chlamydia_OMP; 1.
DR ProDom; PD001717; Chlamydia_OMP; 1.
KW Outer membrane; Transmembrane; Porin.
FT NON_TER 1 1
FT NON_TER 333 333
SQ SEQUENCE 333 AA; 35811 MW; 20460451 2C4C3B3F CRC64;

Query Match 100.0%; Score 73; DB 1; Length 333;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQINKFKSRKACG 13
||||||||||||
Db 298 CQINKFKSRKACG 310

RESULT 2

```
OM1N_CHLPN
ID   OM1N_CHLPN     STANDARD;      PRT;   389 AA.
AC   Q07430;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N16;
RX   MEDLINE=94103736; PubMed=8277245;
RA   Storey C., Lusher M., Yates P., Richmond S.;
RT   "Evidence for Chlamydia pneumoniae of non-human origin.";
RL   J. Gen. Microbiol. 139:2621-2626(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L04982; AAA17397.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     23       BY SIMILARITY.
FT   CHAIN        24    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41628 MW;  801622F05D841967 CRC64;


  Query Match             100.0%;  Score 73;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CQINKFKSRKACG 13
          |||||||||||||
Db    342 CQINKFKSRKACG 354


RESULT   3
OMP1_CHLPN
ID   OMP1_CHLPN     STANDARD;      PRT;   389 AA.
AC   P27455; Q9JQF6;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1 OR CPN0695 OR CP0051.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IOL-207;
RX   MEDLINE=91237311; PubMed=2033374;
RA   Carter M.W., Al-Mahdawi S.A.H., Giles I.G., Treharne J.D.,
RA   Ward M.E., Clarke I.N.;
RT   "Nucleotide sequence and taxonomic value of the major outer membrane
RT   protein gene of Chlamydia pneumoniae IOL-207.";
RL   J. Gen. Microbiol. 137:465-475(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TWAR;
RX   MEDLINE=91244474; PubMed=1840574;
RA   Perez Melgosa M., Kuo C.-C., Campbell L.A.;
RT   "Sequence analysis of the major outer membrane protein gene of
RT   Chlamydia pneumoniae.";
RL   Infect. Immun. 59:2195-2199(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Mitchell W.M., Tharp A.C., Stratton C.W., Sriram S.;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20298986; PubMed=10839753;
RA   Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,
RA   Takeuchi H., Nishida J., Shibata K., Fujinaga R., Yoneda H.,
RA   Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,
RA   Ishii K., Shiba T., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of outer membrane protein genes omp and pmp in the whole
RT   genome sequences of Chlamydia pneumoniae isolates from Japan and the
RT   United States.";
RL   J. Infect. Dis. 181 Suppl 3:S524-S527(2000).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M64064; AAA23143.1; -.
DR   EMBL; M69230; AAA73071.1; -.
DR   EMBL; AF131889; AAD22492.1; -.
```

```
DR   EMBL; AE001652; AAD18834.1; -.
DR   EMBL; AE002167; AAF37944.1; -.
DR   EMBL; AP002547; BAA98902.1; -.
DR   EMBL; AB033787; BAA85940.1; -.
DR   PIR; A43587; A43587.
DR   PIR; A49751; A49751.
DR   TIGR; CP0051; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
FT   SIGNAL        1     23
FT   CHAIN        24    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41620 MW;  15D984151E41F8F2 CRC64;


  Query Match             100.0%;  Score 73;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CQINKFKSRKACG 13
          |||||||||||||
Db    342 CQINKFKSRKACG 354


RESULT   4
OM1A_CHLPS
ID   OM1A_CHLPS     STANDARD;      PRT;   389 AA.
AC   P16567;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OVINE ENZOOTIC ABORTION ISOLATE S26/3;
RX   MEDLINE=90128177; PubMed=2612883;
RA   Herring A.J., Tan T.W., Baxter S., Inglis N.F., Dunbar S.;
RT   "Sequence analysis of the major outer membrane protein gene of an
RT   ovine abortion strain of Chlamydia psittaci.";
RL   FEMS Microbiol. Lett. 53:153-158(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BOVINE ABORTION ISOLATE BA1;
RX   MEDLINE=96189695; PubMed=8605581;
RA   Griffiths P.C., Plater J.M., Martin T.C., Hughes S.L.,
RA   Hughes K.J., Hewinson R.G., Dawson M.;
RT   "Epizootic bovine abortion in a dairy herd: characterization of a
RT   Chlamydia psittaci isolate and antibody response.";
RL   Br. Vet. J. 151:683-693(1995).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X51859; CAA36152.1; -.
DR   EMBL; L39020; AAB02850.1; -.
DR   PIR; S08770; MMCWP3.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41883 MW;  741B5A23ACDBB447 CRC64;


  Query Match             79.5%;  Score 58;  DB 1;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.0051;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    343 QINKMKSRKACG 354


RESULT   5
OM1P_CHLPS
ID   OM1P_CHLPS     STANDARD;      PRT;   392 AA.
AC   Q00087;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FPN/PRING;
RX   MEDLINE=94103736; PubMed=8277245;
RA   Storey C., Lusher M., Yates P., Richmond S.;
RT   "Evidence for Chlamydia pneumoniae of non-human origin.";
RL   J. Gen. Microbiol. 139:2621-2626(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61096; CAA43409.1; -.
DR   PIR; A40371; A40371.
DR   PIR; S16137; S16137.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    392       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   392 AA;  42069 MW;  88B3C5D90BBA26DB CRC64;


  Query Match             79.5%;  Score 58;  DB 1;  Length 392;
  Best Local Similarity   91.7%;  Pred. No. 0.0051;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    346 QINKMKSRKACG 357
```

```
RESULT   6
OM1E_CHLPS
ID   OM1E_CHLPS     STANDARD;      PRT;   402 AA.
AC   P10332;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EAE A22/M;
RA   Pickett M.A., Everson S.J., Clarke I.N.;
RT   "Chlamydia psittaci ewe abortion agent: complete nucleotide sequence
RT   of the major outer membrane protein gene.";
RL   FEMS Microbiol. Lett. 55:229-234(1988).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X12647; CAA31177.1; -.
DR   EMBL; M36703; AAA23146.1; -.
DR   PIR; S05954; MMCWPM.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    402       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   402 AA;  43277 MW;  E6CF00D9DF1EE87A CRC64;


  Query Match             79.5%;  Score 58;  DB 1;  Length 402;
  Best Local Similarity   91.7%;  Pred. No. 0.0052;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     356 QINKMKSRKACG 367


RESULT   7
OMP1_CHLMU
ID   OMP1_CHLMU     STANDARD;      PRT;   387 AA.
AC   P75024; Q04063; Q9X718;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
GN   OMPA OR OMP1 OR TC0052.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MOPN;
RX   MEDLINE=92039057; PubMed=1937036;
RA   Fielder T.J., Pal S., Peterson E.M., la Maza L.M.;
RT   "Sequence of the gene encoding the major outer membrane protein of the
RT   mouse pneumonitis biovar of Chlamydia trachomatis.";
RL   Gene 106:137-138(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MOPN;
RX   MEDLINE=94104488; PubMed=8277858;
RA   Zhang Y.X., Fox J.G., Ho Y., Zhang L., Stills H.F., Smith T.F.;
RT   "Comparison of the major outer-membrane protein (MOMP) gene of mouse
RT   pneumonitis (MoPn) and hamster SFPD strains of Chlamydia trachomatis
RT   with other Chlamydia strains.";
RL   Mol. Biol. Evol. 10:1327-1342(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP.BV.MOUSE / NIGG II;
RA   Carter M.W., Giles I., Everson J.S., Clarke I.N.;
RT   "Chlamydia trachomatis mouse biovar: major outer membrane protein
RT   gene.";
RL   (In) Mardh P.A., la Placa M., Ward M. (eds.);
RL   Proceedings of the european society for chlamydia research and the
RL   second international symposium of Uppsala university centre for std
RL   research, pp.38-38, University of Uppsala, Uppsala (1992).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MOPN / NIGG;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K.,
RA   Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C.,
RA   Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA   Salzberg S.L., Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [5]
RP   SEQUENCE OF 37-375 FROM N.A.
RC   STRAIN=MOPN;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE.
CC   -!- MISCELLANEOUS: MOMP IS RESPONSIBLE FOR THE STRUCTURAL INTEGRITY OF
CC       THE EXTRA-CELLULAR INFECTIOUS ELEMENTARY BODY & THE DEVELOPMENTAL
CC       CONVERSION TO THE PLASTIC AND FRAGILE INTRACELLULAR RETICULATE
CC       BODY.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M64171; AAA23144.1; -.
DR   EMBL; U60196; AAB07068.1; -.
DR   EMBL; X63409; CAA45006.1; -.
DR   EMBL; AE002272; AAF38941.1; -.
DR   EMBL; M73044; AAD29101.1; -.
DR   TIGR; TC0052; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
```

```
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    387       MAJOR OUTER MEMBRANE PROTEIN.
FT   CONFLICT    118    118       F -> Y (IN REF. 5).
FT   CONFLICT    123    123       Y -> F (IN REF. 5).
FT   CONFLICT    198    198       L -> F (IN REF. 1).
FT   CONFLICT    204    204       A -> P (IN REF. 1).
SQ   SEQUENCE   387 AA;  42009 MW;  4FD6FDC23248E0A2 CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 387;
  Best Local Similarity   75.0%;  Pred. No. 0.038;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| |||| :||
Db    341 QLNKMKSRKSCG 352


RESULT   8
OM1D_CHLTR
ID   OM1D_CHLTR     STANDARD;      PRT;   393 AA.
AC   Q46409;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN, SEROVAR D PRECURSOR (MOMP).
GN   OMPA OR OMP1 OR CT681.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/B-120;
RX   MEDLINE=93013014; PubMed=1398119;
RA   Sayada C., Denamur E., Elion J.;
RT   "Complete sequence of the major outer membrane protein-encoding gene
RT   of Chlamydia trachomatis serovar Da.";
RL   Gene 120:129-130(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/IU-71960;
RX   MEDLINE=98339860; PubMed=9673241;
RA   Stothard D.R., Boguslawski G., Jones R.B.;
RT   "Phylogenetic analysis of the Chlamydia trachomatis major outer
RT   membrane protein and examination of potential pathogenic
RT   determinants.";
RL   Infect. Immun. 66:3618-3625(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/CX;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62918; CAA44701.1; -.
DR   EMBL; AF063195; AAC31436.2; -.
DR   EMBL; AE001338; AAC68276.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    393       MAJOR OUTER MEMBRANE PROTEIN, SEROVAR D.
SQ   SEQUENCE   393 AA;  42438 MW;  8CD692FD3EFF21D6 CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 393;
  Best Local Similarity   75.0%;  Pred. No. 0.039;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| |||| :||
Db    347 QLNKMKSRKSCG 358


RESULT   9
OM1E_CHLTR
ID   OM1E_CHLTR     STANDARD;      PRT;   393 AA.
AC   P17451;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN, SEROVAR E PRECURSOR (MOMP).
GN   OMPA OR OMP1E.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BOUR / SEROVAR E;
RX   MEDLINE=90287737; PubMed=2356137;
RA   Peterson E.M., Markoff B.A., de la Maza L.M.;
RT   "The major outer membrane protein nucleotide sequence of Chlamydia
RT   trachomatis, serovar E.";
RL   Nucleic Acids Res. 18:3414-3414(1990).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52557; CAA36791.1; -.
DR   PIR; S10201; MMCWTE.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    393       MAJOR OUTER MEMBRANE PROTEIN, SEROVAR E.
SQ   SEQUENCE   393 AA;  42424 MW;  AB2B82D16027B361 CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 393;
  Best Local Similarity   75.0%;  Pred. No. 0.039;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
```

```
          |:|| ||||:||
Db    347 QLNKMKSRKSCG 358


RESULT  10
OM1L_CHLTR
ID   OM1L_CHLTR     STANDARD;      PRT;   393 AA.
AC   P19542;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN, SEROVAR L1 PRECURSOR (MOMP).
GN   OMPA OR OMP1L1.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pickett M.A., Ward M.E., Clarke I.N.;
RT   "Complete nucleotide sequence of the major outer membrane protein
RT   gene from Chlamydia trachomatis serovar L1.";
RL   FEMS Microbiol. Lett. 42:185-190(1987).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M36533; AAA23142.1; -.
DR   PIR; S06259; S06259.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    393       MAJOR OUTER MEMBRANE PROTEIN, SEROVAR L1.
SQ   SEQUENCE   393 AA;  42543 MW;  7A952839408EE2DF CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 393;
  Best Local Similarity   75.0%;  Pred. No. 0.039;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| ||||:||
Db    347 QLNKMKSRKSCG 358


RESULT  11
OM1B_CHLTR
ID   OM1B_CHLTR     STANDARD;      PRT;   394 AA.
AC   P23421;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN, SEROVAR B PRECURSOR (MOMP).
GN   OMPA OR OMP1B.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87307955; PubMed=3040664;
RA   Stephens R.S., Sanchez-Pescador R., Wagar E.A., Inouye C., Urdea M.S.;
RT   "Diversity of Chlamydia trachomatis major outer membrane protein
RT   genes.";
RL   J. Bacteriol. 169:3879-3885(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B/Tw-5/OT;
RA   Dean D.A.;
RT   "Sequence analysis of the major outer membrane protein gene (ompA) of
RT   Chlamydia trachomatis.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M17342; -; NOT_ANNOTATED_CDS.
DR   EMBL; AF304856; AAG41414.1; -.
DR   PIR; S11010; MMCWTB.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    394       MAJOR OUTER MEMBRANE PROTEIN, SEROVAR B.
SQ   SEQUENCE   394 AA;  42528 MW;  C364233145A69301 CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 394;
  Best Local Similarity   75.0%;  Pred. No. 0.039;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| ||||:||
Db    348 QLNKMKSRKSCG 359


RESULT  12
OM1M_CHLTR
ID   OM1M_CHLTR     STANDARD;      PRT;   394 AA.
AC   P06597;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN, SEROVAR L2 PRECURSOR (MOMP).
GN   OMPA OR OMP1L2.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87057033; PubMed=2946665;
RA   Stephens R.S., Mullenbach G., Sanchez-Pescador R., Agabian N.;
RT   "Sequence analysis of the major outer membrane protein gene from
RT   Chlamydia trachomatis serovar L2.";
RL   J. Bacteriol. 168:1277-1282(1986).
RN   [2]
RP   SEQUENCE OF 23-27.
RC   STRAIN=L2/434/BU;
RA   Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA   Comanducci M., Christianen G., Birkelund S., Vtretou E., Ratti G.,
RA   Pallini V.;
```

```
RL   Submitted (SEP-1994) to the SWISS-PROT data bank.
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14738; AAA23151.1; -.
DR   PIR; S11012; S11012.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    394       MAJOR OUTER MEMBRANE PROTEIN, SEROVAR L2.
SQ   SEQUENCE   394 AA;  42550 MW;  BB5B7B80EB289CA5 CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 394;
  Best Local Similarity   75.0%;  Pred. No. 0.039;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| ||||:||
Db    348 QLNKMKSRKSCG 359


RESULT  13
OM1F_CHLTR
ID   OM1F_CHLTR     STANDARD;      PRT;   395 AA.
AC   P16155;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN, SEROVAR F PRECURSOR (MOMP).
GN   OMPA OR OMP1F.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IC-CAL3 / SEROVAR F;
RX   MEDLINE=90192102; PubMed=2315025;
RA   Zhang Y.X., Morrison S.G., Caldwell H.D.;
RT   "The nucleotide sequence of major outer membrane protein gene of
RT   Chlamydia trachomatis serovar F.";
RL   Nucleic Acids Res. 18:1061-1061(1990).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52080; CAA36299.1; -.
DR   PIR; S08463; MMCWTF.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    395       MAJOR OUTER MEMBRANE PROTEIN, SEROVAR F.
SQ   SEQUENCE   395 AA;  42586 MW;  7F90FFDEEC264ACF CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 395;
  Best Local Similarity   75.0%;  Pred. No. 0.039;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |:|| ||||:||
Db    349 QLNKMKSRKSCG 360


RESULT  14
OM1A_CHLTR
ID   OM1A_CHLTR     STANDARD;      PRT;   396 AA.
AC   P23732;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN, SEROVAR A PRECURSOR (MOMP).
GN   OMPA OR OMP1A.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SA1/OT / SEROVAR A;
RX   MEDLINE=91045088; PubMed=2235504;
RA   Hayes L.J., Clarke I.N.;
RT   "Nucleotide sequence of the major outer membrane protein gene of
RT   Chlamydia trachomatis strain A/SA1/OT.";
RL   Nucleic Acids Res. 18:6136-6136(1990).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M58938; AAA23141.1; -.
DR   EMBL; M33635; AAA92785.1; -.
DR   PIR; S12799; S12799.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    396       MAJOR OUTER MEMBRANE PROTEIN, SEROVAR A.
SQ   SEQUENCE   396 AA;  42877 MW;  2F9D3B0CE2D08162 CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 396;
  Best Local Similarity   75.0%;  Pred. No. 0.039;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
```

```
          |:|| ||||:||
Db    350 QLNKMKSRKSCG 361


RESULT  15
OM1C_CHLTR
ID   OM1C_CHLTR     STANDARD;      PRT;   397 AA.
AC   P08780;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN, SEROVAR C PRECURSOR (MOMP).
GN   OMPA OR OMP1 OR OMP1C.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87307955; PubMed=3040664;
RA   Stephens R.S., Sanchez-Pescador R., Wagar E.A., Inouye C., Urdea M.S.;
RT   "Diversity of Chlamydia trachomatis major outer membrane protein
RT   genes.";
RL   J. Bacteriol. 169:3879-3885(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C/TW3;
RX   MEDLINE=20407420; PubMed=10950788;
RA   Dean D., Suchland R.J., Stamm W.E.;
RT   "Evidence for long-term cervical persistence of Chlamydia trachomatis
RT   by omp1 genotyping.";
RL   J. Infect. Dis. 182:909-916(2000).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL OMP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M17343; AAA23156.1; -.
DR   EMBL; AF202455; AAG09443.1; -.
DR   PIR; S11011; MMCWTC.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    397       MAJOR OUTER MEMBRANE PROTEIN, SEROVAR C.
SQ   SEQUENCE   397 AA;  42892 MW;  0047BCDB108E5309 CRC64;


  Query Match             72.6%;  Score 53;  DB 1;  Length 397;
  Best Local Similarity   75.0%;  Pred. No. 0.039;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |:|| ||||:||
Db     351 QLNKMKSRKSCG 362


Search completed: March 26, 2002, 13:40:46
Job time: 260 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 26, 2002, 13:40:15 ; Search time 79.01 Seconds
(without alignments)
24.067 Million cell updates/sec

Title: US-09-709-201-101
Perfect score: 73
Sequence: 1 CQINKFKSRKACG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73	100.0	389	2	Q08085	Q08085 chlamydia p
2	67	91.8	336	2	Q9XBF5	Q9xbf5 chlamydophi
3	58	79.5	326	2	Q9K5C5	Q9k5c5 chlamydia p
4	58	79.5	330	2	Q9XBF3	Q9xbf3 chlamydophi
5	58	79.5	337	2	Q9XBF6	Q9xbf6 chlamydophi
6	58	79.5	340	2	Q9XBF2	Q9xbf2 chlamydophi
7	58	79.5	341	2	Q9X717	Q9x717 chlamydophi
8	58	79.5	380	2	Q9AII1	Q9aii1 chlamydophi
9	58	79.5	381	2	Q9AII2	Q9aii2 chlamydophi
10	58	79.5	382	2	Q9AIJ9	Q9aij9 chlamydia p
11	58	79.5	388	2	Q9AIK1	Q9aik1 chlamydia p
12	58	79.5	388	2	Q9AIK0	Q9aik0 chlamydia p
13	58	79.5	389	2	Q9APM4	Q9apm4 chlamydophi
14	58	79.5	389	2	Q9AIH9	Q9aih9 chlamydophi
15	58	79.5	390	2	Q9AIJ5	Q9aij5 chlamydia p
16	58	79.5	391	2	Q46235	Q46235 chlamydia p
17	58	79.5	391	2	Q9AIJ2	Q9aij2 chlamydia p
18	58	79.5	392	2	Q9AIJ4	Q9aij4 chlamydia p
19	58	79.5	392	2	Q99QB0	Q99qb0 chlamydophi
20	58	79.5	395	2	Q9AIJ7	Q9aij7 chlamydia p
21	58	79.5	397	2	Q9AIJ8	Q9aij8 chlamydia p
22	58	79.5	402	2	Q46203	Q46203 chlamydia p
23	58	79.5	402	2	Q46236	Q46236 chlamydia p
24	58	79.5	402	2	Q46193	Q46193 chlamydia p
25	58	79.5	402	2	Q9AII0	Q9aii0 chlamydia p
26	55	75.3	322	2	Q9XBF1	Q9xbf1 chlamydia s
27	55	75.3	376	2	Q9AII9	Q9aii9 chlamydia s
28	55	75.3	385	2	Q9AII7	Q9aii7 chlamydia s
29	55	75.3	385	2	Q9AII6	Q9aii6 chlamydia s
30	55	75.3	386	2	Q9AII5	Q9aii5 chlamydia s
31	55	75.3	387	2	Q9AIJ1	Q9aij1 chlamydia s
32	55	75.3	387	2	Q9AIJ0	Q9aij0 chlamydia s
33	55	75.3	389	2	Q9AII4	Q9aii4 chlamydia s
34	55	75.3	391	2	Q9AII3	Q9aii3 chlamydia s
35	55	75.3	396	2	Q9AII8	Q9aii8 chlamydia s
36	55	75.3	402	2	Q9AIJ6	Q9aij6 chlamydia p
37	55	75.3	402	2	Q9AIJ3	Q9aij3 chlamydia p
38	53	72.6	103	2	Q9S6E6	Q9s6e6 chlamydia t
39	53	72.6	106	2	Q9S6E9	Q9s6e9 chlamydia t
40	53	72.6	107	2	Q9S6E7	Q9s6e7 chlamydia t
41	53	72.6	108	2	Q9X4V8	Q9x4v8 chlamydia t
42	53	72.6	108	2	Q9S6F2	Q9s6f2 chlamydia t
43	53	72.6	108	2	Q9S6E8	Q9s6e8 chlamydia t
44	53	72.6	108	2	Q9S6E1	Q9s6e1 chlamydia t
45	53	72.6	109	2	Q9ZFE8	Q9zfe8 chlamydia t

ALIGNMENTS

```
RESULT   1
Q08085
ID   Q08085      PRELIMINARY;      PRT;   389 AA.
AC   Q08085;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (MOMP).
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KOALA TYPE 1;
RX   MEDLINE=94171025; PubMed=8125292;
RA   Girjes A.A., Carrick F.N., Lavin M.F.;
RT   "Remarkable sequence relatedness in the DNA encoding the major outer
RT   membrane protein of Chlamydia psittaci (koala type I) and Chlamydia
RT   pneumoniae.".
RL   Gene 138:139-142(1994).
CC   -!- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
CC       BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
CC       THE INTRACELLULAR RETICULATE BODY MEMBRANE.
CC   -!- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
CC       MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
DR   EMBL; X72023; CAA50906.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   PRINTS; PR01334; CHLAMIDIAOMP.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
KW   Outer membrane; Transmembrane; Porin; Signal.
FT   SIGNAL        1     23       BY SIMILARITY.
FT   CHAIN        24    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41579 MW;  5DC50E85A6F4E50F CRC64;


  Query Match          100.0%;  Score 73;  DB 2;  Length 389;
  Best Local Similarity  100.0%;  Pred. No. 8e-06;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CQINKFKSRKACG 13
```

```
                  ||||||||||||
Db     342 CQINKFKSRKACG 354


RESULT   2
Q9XBF5
ID   Q9XBF5       PRELIMINARY;      PRT;   336 AA.
AC   Q9XBF5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER MEMBRANE PROTEIN (FRAGMENT).
OS   Chlamydophila pecorum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=85991;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=66P130;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
DR   EMBL; M73034; AAD38209.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   NON_TER     336    336
SQ   SEQUENCE   336 AA;  36698 MW;  0AC8C33EB3E0A61A CRC64;


  Query Match             91.8%;  Score 67;  DB 2;  Length 336;
  Best Local Similarity   92.3%;  Pred. No. 9.5e-05;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CQINKFKSRKACG 13
           ||||| |||||||
Db     301 CQINKLKSRKACG 313


RESULT   3
Q9K5C5
ID   Q9K5C5       PRELIMINARY;      PRT;   326 AA.
AC   Q9K5C5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER MEMBRANE PROTEIN 1 (FRAGMENT).
GN   OMP1.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R54;
RX   PubMed=10919838;
RA   Herrmann B., Rahman R., Bergstroem S., Bonnedahl J., Olsen B.;
RT   "Chlamydophila abortus in a Brown Skua (Catharacta antarctica
RT   lonnbergi) from a Subantarctic Island.";
RL   Appl. Environ. Microbiol. 66:3654-3656(2000).
DR   EMBL; AJ243525; CAB96859.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   PRINTS; PR01334; CHLAMIDIAOMP.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   NON_TER     326    326
SQ   SEQUENCE   326 AA;  35345 MW;  6C5A20C8913743C8 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 326;
  Best Local Similarity   91.7%;  Pred. No. 0.0046;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     290 QINKMKSRKACG 301


RESULT   4
Q9XBF3
ID   Q9XBF3       PRELIMINARY;      PRT;   330 AA.
AC   Q9XBF3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER MEMBRANE PROTEIN (FRAGMENT).
OS   Chlamydophila pecorum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=85991;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L71;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
DR   EMBL; M73039; AAD38211.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   NON_TER     330    330
SQ   SEQUENCE   330 AA;  35834 MW;  FCB6D9332991BB09 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 330;
  Best Local Similarity   91.7%;  Pred. No. 0.0046;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     296 QINKLKSRKACG 307


RESULT   5
Q9XBF6
ID   Q9XBF6       PRELIMINARY;      PRT;   337 AA.
AC   Q9XBF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER MEMBRANE PROTEIN (FRAGMENT).
OS   Chlamydophila pecorum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=85991;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1710S;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
DR   EMBL; M73033; AAD38208.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   NON_TER     337    337
SQ   SEQUENCE   337 AA;  36817 MW;  A173FB2597C84E6B CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 337;
```

```
  Best Local Similarity   91.7%;  Pred. No. 0.0047;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     303 QINKLKSRKACG 314


RESULT   6
Q9XBF2
ID   Q9XBF2      PRELIMINARY;      PRT;   340 AA.
AC   Q9XBF2;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER MEMBRANE PROTEIN (FRAGMENT).
OS   Chlamydophila pecorum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=85991;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LW613;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
DR   EMBL; M73042; AAD38212.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   NON_TER     340    340
SQ   SEQUENCE   340 AA;  36968 MW;  F571E822D1DE0AA3 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 340;
  Best Local Similarity   91.7%;  Pred. No. 0.0047;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     306 QINKLKSRKACG 317


RESULT   7
Q9X717
ID   Q9X717      PRELIMINARY;      PRT;   341 AA.
AC   Q9X717;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydophila abortus.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LW508;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
DR   EMBL; M73040; AAD29103.1; -.
DR   InterPro; IPR000604; Chlamydia_OMP.
DR   Pfam; PF01308; Chlamydia_OMP; 1.
DR   ProDom; PD001717; Chlamydia_OMP; 1.
FT   NON_TER       1      1
FT   NON_TER     341    341
SQ   SEQUENCE   341 AA;  36762 MW;  B5933C9BF6AAF171 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 341;
  Best Local Similarity   91.7%;  Pred. No. 0.0048;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     307 QINKMKSRKACG 318


RESULT   8
Q9AII1
ID   Q9AII1      PRELIMINARY;      PRT;   380 AA.
AC   Q9AII1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydophila pecorum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=85991;
RN   [1]
RP   SEQUENCE OF 40-352 FROM N.A.
RC   STRAIN=L71;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L71;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L71;
RA   Everett K.D.E., Hambly W.A., Andersen A.A.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF269280; AAK00261.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     15       POTENTIAL.
FT   CHAIN        16    380       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   380 AA;  41047 MW;  829A18D3C5A85008 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 380;
  Best Local Similarity   91.7%;  Pred. No. 0.0052;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 QINKFKSRKACG 13
           |||| |||||||
Db     334 QINKLKSRKACG 345


RESULT   9
Q9AII2
ID   Q9AII2      PRELIMINARY;      PRT;   381 AA.
AC   Q9AII2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydophila pecorum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=85991;
RN   [1]
RP   SEQUENCE OF 34-369 FROM N.A.
```

```
RC   STRAIN=1710S;
RX   MEDLINE=93123168; PubMed=8419295;
RA   Kaltenboeck B., Kousoulas K.G., Storz J.;
RT   "Structures of and allelic diversity and relationships among the major
RT   outer membrane protein (ompA) genes of the four chlamydial species.";
RL   J. Bacteriol. 175:487-502(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1710S;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1710S;
RA   Everett K.D.E., Hambly W.A., Andersen A.A.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF269279; AAK00260.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     15       POTENTIAL.
FT   CHAIN        16    381       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   381 AA;  41332 MW;  29406725CF9D3512 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 381;
  Best Local Similarity   91.7%;  Pred. No. 0.0052;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    335 QINKLKSRKACG 346


RESULT  10
Q9AIJ9
ID   Q9AIJ9      PRELIMINARY;      PRT;   382 AA.
AC   Q9AIJ9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MENINGOPNEUMONITIS, MN, ATCC VR122;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269262; AAK00243.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1      2       POTENTIAL.
FT   CHAIN         3    382       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   382 AA;  41231 MW;  6917171A9A69303B CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 382;
  Best Local Similarity   91.7%;  Pred. No. 0.0053;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    336 QINKMKSRKACG 347


RESULT  11
Q9AIK1
ID   Q9AIK1      PRELIMINARY;      PRT;   388 AA.
AC   Q9AIK1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VS225;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269259; AAK00240.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     19       POTENTIAL.
FT   CHAIN        20    388       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   388 AA;  41573 MW;  8E232D22C9B9948D CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 388;
  Best Local Similarity   91.7%;  Pred. No. 0.0053;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    342 QINKMKSRKACG 353


RESULT  12
Q9AIK0
ID   Q9AIK0      PRELIMINARY;      PRT;   388 AA.
AC   Q9AIK0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CALIFORNIA TURKEY 1, CT1;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269260; AAK00241.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     19       POTENTIAL.
FT   CHAIN        20    388       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   388 AA;  42053 MW;  96E675B3F69F708B CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 388;
  Best Local Similarity   91.7%;  Pred. No. 0.0053;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    342 QINKMKSRKACG 353


RESULT  13
Q9APM4
ID   Q9APM4      PRELIMINARY;      PRT;   389 AA.
```

```
AC   Q9APM4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN   OMP1.
OS   Chlamydophila abortus.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LLG;
RX   MEDLINE=20569239; PubMed=11119563;
RA   Vretou E., Psarrou E., Kaisar M., Vlisidou I., Salti-Montesanto V.,
RA   Longbottom D.;
RT   "Identification of protective epitopes by sequencing of the major
RT   outer membrane protein gene of a variant strain of Chlamydia psittaci
RT   serotype 1.";
RL   Infect. Immun. 69:607-612(2001).
DR   EMBL; AF272945; AAG53881.1; -.
KW   Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41897 MW;  20513C69C7DBAAF5 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.0053;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    343 QINKMKSRKACG 354


RESULT  14
Q9AIH9
ID   Q9AIH9      PRELIMINARY;      PRT;   389 AA.
AC   Q9AIH9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR.
GN   OMPA.
OS   Chlamydophila caviae.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83557;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GUINEA PIG INCLUSION CONJUNCTIVITIS, GPIC, ATCC VR813;
RX   MEDLINE=89212917; PubMed=2707861;
RA   Zhang Y.X., Morrison S.G., Caldwell H.D., Baehr W.;
RT   "Cloning and sequence analysis of the major outer membrane protein
RT   genes of two Chlamydia psittaci strains.";
RL   Infect. Immun. 57:1621-1625(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GUINEA PIG INCLUSION CONJUNCTIVITIS, GPIC, ATCC VR813;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269282; AAK00263.1; -.
KW   Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    389       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   389 AA;  41932 MW;  2527A820C76F8310 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.0053;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    343 QINKMKSRKACG 354


RESULT  15
Q9AIJ5
ID   Q9AIJ5      PRELIMINARY;      PRT;   390 AA.
AC   Q9AIJ5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN PRECURSOR (FRAGMENT).
GN   OMPA.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW JERSEY 1, NJ1;
RX   MEDLINE=21078680; PubMed=11211261;
RA   Bush R.M., Everett K.D.;
RT   "Molecular evolution of the Chlamydiaceae.";
RL   Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR   EMBL; AF269266; AAK00247.1; -.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     20       POTENTIAL.
FT   CHAIN        21    390       MAJOR OUTER MEMBRANE PROTEIN.
SQ   SEQUENCE   390 AA;  42042 MW;  B62858403DBFA4E6 CRC64;


  Query Match             79.5%;  Score 58;  DB 2;  Length 390;
  Best Local Similarity   91.7%;  Pred. No. 0.0054;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 QINKFKSRKACG 13
          |||| |||||||
Db    344 QINKMKSRKACG 355


Search completed: March 26, 2002, 13:40:16
Job time: 230 sec
```